PHD

Essays on lending cycles, bank risk-taking and monetary policy

Prat, Marion

Award date:
2016

Awarding institution:
University of Bath

Link to publication

Essays on lending cycles, bank risk-taking and monetary policy

submitted by

Marion Prat

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Economics

September 2015

Acknowledgments

This thesis owes a lot to the patience and understanding provided by my supervisors, Chris Martin, Christos Ioannidis and Matteo de Tina. I specially thank Chris who read most of it and never seemed to doubt its worth. My internal and external examiners, Alessandro Marchesiani and George Bratsiosis, made helpful comments and rendered the last stage of my thesis remarkably bearable. A huge thank you to my colleagues and friends in the Department of Economics, and in particular to Bingsong Wang, Jose Fernandez, Alaa Al khourdajie, Lory Barile, Ola Aworinde and Elvis Moleka, who bore equally my moments of exaltation and desperation. A special thank you to Nikos Kokonas for always showing an interest in my research and more importantly talking about his. I would also like to thank the administrative staff for taking such a good care of me and particularly Sally Rose, Catherine Adams and Emma Mulhern. I would not have been able to do much without the love and support of my family, and in particular my parents Regine and Alain, my grandmothers Fernande and Monique as well as of Mathieu, Julie, Dominique and Christian. Finally, Sabine Graf, Jamie O'Neill and Yvonne Van Rossenberg were the best housemates I could ever dream of.

Abstract

The thesis is constructed around two themes. In the first part, we review the main workhorse macroeconomic models emphasizing a role for financial frictions in business cycle fluctuations. Their main contributions was to develop the two concepts of the external finance premium and asset collateral value (Bernanke and Gertler, 1989; Kiyotaki and Moore, 1997) and to theorize the organization of the financial sector, therefore providing new insights into the nature of intermediation spreads and capital solvency ratios (Holmstrom and Tirole, 1997; Gertler and Kiyotaki, 2011). We investigate the interaction between inflation dynamics and the financial accelerator effect in the model Bernanke, Gertler and Gilchrist (1999) and show that the magnitude of the financial accelerator effect depends crucially on the slope of the Phillips curve. In particular, the flatter the Philips curve, the stronger is the financial accelerator effect.

In the second, empirical part, we first investigate the risk-taking channel of monetary policy using US bank lending survey data over the 17 years that preceded the 2007-2009 financial crisis. We find robust evidence that an expansionary monetary policy led to an immediate tightening of bank lending standards followed by a subsequent loosening as predicted by recent theories of bank risk-taking. In addition, the term spread may play a pivotal role in this transmission mechanism. Finally, we empirically investigate the hypothesis that financial conditions critically affect the ability of commercial banks to hedge liquidity shocks. Using US commercial banks' balance sheet data over the period 1990-2007, we show that the CP-Tbill spread can help identify regimes where this hedge is effective in protecting bank loan supply from adverse financial market conditions using nonlinear methods.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The 2007-2009 financial and banking crisis that started in the US and spred globally disrupted for a long time the level of activity and the way monetary policy is conducted. It also hihlighted the need to better understand and document the links between monetary policy and financial fragility, taking into account their highly nonlinear nature.

Large spikes in spreads and volatility indices are a key feature of financial crises. One thread running through the thesis is an examination of the different components of the risk premia affecting the investment decisions of economic agents. On one side financial accelerator models postulate the existence of an external finance premium originating in problems of asymmetric information between borrowers and lenders (Bernanke and Gertler, 1989). On the other side, theories of the risk-taking channels of monetary policy are consistent with time-varying risk premia which may be affected by monetary policy (Bekaert et. al., 2013) and by the relative share of financial intermediaries or of certain types of creditors in the economy (Adrian and Shin, 2011).

Our first contribution (chapters 2,3 and 4) is theoretical and examines one of the determinants of the financial accelerator effect as an amplification mechanism. The idea that adverse developments in the real economy may lead to serious disruptions in the financial sector, which in turn serves to amplify adverse macroeconomic conditions goes back to Fisher (1933) and Bernanke (1983) who analyze the mechanisms underlying the Great Depression that hit the United States in the 1930s. Following theoretical advances in the literature on asymmetric information problems, the interaction between financial frictions and output dynamics were introduced in general equilibrium models in the 1980s leading to the seminal papers of Bernanke and Gertler (1989) and Kiyotaki and Moore (1997).

Financial accelerator models can be summarized by two key concepts, the notion of external finance premium and the notion of collateral value of assets. The first refers to the idea that firms' (borrowers) inside return to funds is higher than the outside return to funds or the opportunity cost of funds. This idea is emphasized by Bernanke and Gertler (1989) and more generally by the models based on the costly state verification approach. The second was emphasized by Kiyotaki and Moore (1997) who model the most explicitly the interplay between asset price dynamics and credit constraints.

Bernanke, Gertler and Gilchrist (1999) adapted the costly state verification framework to a quantitative general equilibrium model and evaluate the magnitude of the financial accelerator effect following a contractionary monetary shock. The non-neutrality of money result relies on the assumption that prices do not adjust instantaneously to changes in the economic environment. In particular, it is assumed that only a fraction of the prices are adjusted in each period (Calvo model). Monopolistic firms incorporate this feature into their profit-maximizing problem, the solution of which, aggregated over all firms, yields the Phillips curve. It says that in equilibrium current inflation depends on the current real marginal cost and expected future inflation. The higher the degree of price stickiness, the lower the sensitivity of current inflation to real marginal costs and the flatter the Phillips curve. We show that this parameter also determines the strength of the financial accelerator effect.

Our second contribution (chapter 5) relates to the risk-taking channel of monetary policy. Borio and Zhu (2012) defined this notion as "the impact of changes in policy rates on either risk perceptions or risk tolerance, and hence on the degree of risk in portfolios, on the pricing of assets and on the price and non-price terms of the extension of funding". It has been linked to time-varying risk premia (Bekaert et. al., 2013), the procyclicality of bank leverage (Adrian and Shin, 2011) and the existence of a global financial financial cycle (Bruno and Shin, 2015; Miranda Agrippino and Rey, 2013). It therefore raises important policy issues ranging from the interconnection between macroprudential and microprudential issues and the regulation of the international monetary system.

On the empirical side, there is significant evidence in favour of the risk-taking channel of monetary policy using micro-data (Jimenez et. al. 2014; Ioannidou et. al. 2015) and relatively little evidence using aggregate data (Buch et. al., 2014). Time series evidence is needed to evaluate the economic significance of this mechanism as well to provide a sense of the timing of these effects. Although cross-sectional evidence emphasizes an interpretation of the risk-taking channel closely linked to the existence of moral hazard problems within the financial sector, time series evidence must rely on the information contained in an extra dimension of the extension of credit well summarized by bank lending standards. Bank lending standards represent one dimension of bank risk-taking as it is a non-price term of the loan contracts extended by banks. On the theoretical side, Dell'Ariccia et. al. (2014) model the different channels through which changes in real risk-free rates may affect simultaneously intermediation spreads, the extension of credit and bank monitoring. Taking this model as our point of reference, our objective of this study is to examine the dynamic interactions between changes in the Fed Funds rate and bank lending standards. Can we observe systematic patterns of causality in the relationship between the policy rate and bank lending standards in the pre 2007-2009 crisis period in the US?

We find robust evidence that an expansionary monetary policy in the US over the 15 years that preceded the 2007-2009 financial crisis led to an immediate tightening of bank lending standards followed by a subsequent loosening. This is consistent with an immediate risk-shifting effect of monetary policy giving way to a risk-taking effect in the medium term. We cannot however conclude, within the empirical framework adopted in the chapter and due to data limitations, whether the loosening of lending standards is due to increased bank risk-taking or a higher demand for loans by riskier borrowers. This relationship is also economically significant. In addition, we find consistent evidence linking falls in the Fed Funds rate, to a widening of the term spread and the easing of credit conditions for businesses. In view of the significant predictive power of bank lending standards for real activity, these results may also be linked to the well-known ability of the term spread to predict GDP (Adrian and Shin, 2011). Our results also suggest that demand side factors and firm balance sheet effects should not be neglected when evaluating the importance of changes in the policy rate on bank risk-taking.

Our third contribution (chapter 6), we use nonlinear methods to investigate the dual role of commercial banks in the supply of credit depending on the state of the business cycle. Employing a threshold VAR methodology, we investigate the presence and the role of nonlinear relationships between changes in the CP-Tbill spread and commercial banks' balance sheet variables in the US in the run-up the the 2007-2009 financial crisis. The paper-bill has been extensively used in the literature on the financial accelerator as a gauge of the extent of frictions in financial markets. In the banking literature, increases in the spread have been shown to signal the simultaneous drawdown of credit lines by borrowers and the re-intermediation of funds into bank deposits by risk-averse investors (Gatev and Strahan, 2006).

We make several contributions to the literature. We are the first to formally test and find significant evidence in favour of threshold linearity with respect to the level of the paper-bill spread in the relationship between the latter and bank balance sheet variables. Moreover, the estimated threshold is consistent across a range of models including both asset and liability side variables and across categories of banks (large, small and foreign). It clearly differentiates a low spread regime characterized by low volatility and broadly corresponding to periods of low policy rates.

Second, we show that commercial banks are more likely to reduce credit supply in the low spread regime following adverse shocks to financial conditions where we observe sharp increases in commercial and industrial loans accompanied by sharp withradawals of deposits and falls in real estate loans and liquid assets. On the contrary, in the high spread regime, deposits and liquid assets remain stable or increase, allowing banks to hedge the liquidity risk exposure due to the unexpected drawdown of credit lines. Our findings suggest that periods of very low paper-bill spread may lead to a disintermediation of funds that may have significant effects on the ability of commercial banks to respond to financial disruptions.

Chapter 2

Financial accelerator models: borrowers' balance sheet

2.1 Introduction

In this chapter, we discuss in some detail the "old" generation of financial accelerator models developed by Bernanke and Gertler (1989), Carlstrom and Fuerst (1997), Bernanke et. al. (1999) and Kiyotaki and Moore (1997). These models build on the concepts developed in the literature on contracts and incentives to show how financial frictions arising at the microeconomic level can have sizable macroeconomic effects. Asymmetric information problems between lenders and borrowers lead to the failure of the Modigliani-Miller theorem and transform borrowers' net worth into a key state variable, constraining the leverage of borrowers in equilibrium. Shocks that affect the strength of this constraint will therefore be amplified by this mechanism, generating a "financial accelerator" effect.

Financial accelerator models can be summarized by two key concepts, the notion of external finance premium and the notion of collateral value of assets. The first refers to the idea that firms' (borrowers) inside return to funds is higher than the outside return to funds or the opportunity cost of funds. It is this idea that is emphasized by Bernanke and Gertler (1989) and more generally by the models based on the costly state verification approach. The second was emphasized by Kiyotaki and Moore (1997) who model the most explicitly the interplay between asset price dynamics and credit constraints. Depending on whether the focus is placed on one or the other, the key price variable will be an interest rate spread or capital asset prices. The different modelling choices employed in these models capture different types of inefficiencies at the macroeconomic level and imply different propagation mechanisms.

Table 2.1 provides a general overview of the main features of the models. Bernanke and Gertler (1989)[BG], Carlstrom and Fuerst (1997)[CF] and Bernanke, Gertler and Gilchrist (1999)[BGG] introduce the costly state verification framework pioneered by Townsend (1979) into general equilibrium models. This approach is based on the idea that borrowers may have an

Table 2.1: Financial accelerator models

	FC	IC	Risk premium	Asset values	Asymmetric effects
BG	Yes	Yes	Yes for some	No	Yes
KM	Yes	No	No	Yes	No
CF	No	Yes	Yes	No	No
BGG	No	Yes	Yes	Yes	No

FC: Full collateralizarion and IC: Incomplete collateralization

incentive to misreport the outcome of their investment to lenders in the presence of asymmetric information and because of limited liability. In equilibrium, credit market conditions affect the total amount of lending via the cost of external finance. In contrast, Kiyotaki and Moore (1997)[KM] based the financial friction in their model on the notion of the inalienability of human capital developed by Hart and Moore (1994). In equilibrium some agents are credit rationed and asset prices are the key variable defined as the outcome of the microeconomic problem.

The first approach allows for incomplete collateralization so that some lending is risky, whereas the second is characterized by full collateralisation in equilibrium. In the full collateralization case, borrowers are certain to be able to repay loans either because the amount of debt is low or because collateral asset values are high. As result default cannot be an equilibrium outcome. Incomplete collateralization corresponds to the opposite situation. Finally, although BGG allow for a role for asset values, their interaction with the financial accelerator mechanism is not determined as the outcome of the contracting problem between borrowers and lenders as in KM. It follows from the behaviour of another set of agents who interact with borrowers and lenders in the general equilibrium model. As a result, in BGG asset prices are not inefficient as in KM.

Following the seminal work of BG who introduced the costly state verification framework into a neoclassical growth model, CF and BGG developed two general equilibrium models with closed-form solutions to analyse quantitatively how financial frictions affect the real economy. KM developed an asset pricing model in a general equilibrium setting emphasizing the process as in BG rather than the measurement.

2.2 Bernanke and Gertler (1989)

The model developed by Bernanke and Gertler (1989) [BG] achieves several goals. First, using the neoclassical growth framework, it provides strong theoretical foundations to the claim that problems of imperfect information in the credit market can have significant macroeconomic effects through their effect on the determination of the supply of capital (or investment). Second, it shows that one implication of this claim is to create a wedge between the returns to internal and external funds. Third, it establishes that financial frictions are both a source of inefficiency and a source of persistence. Fourth, it shows how these results may account for a debt-deflation phenomenon.

2.2.1 The basic setup

The framework is an overlapping generations model with two periods in which there are two types of agents, households (who are also lenders) and entrepreneurs (who can become borrowers). Both provide labour inelastically in a competitive labour market and are risk-neutral. Their savings in the second period are equal to their labour income minus their consumption in the first period. Their problem is to choose their levels of consumption in the two periods that maximize their expected lifetime utility.

There are two goods, an output good and a capital good. The output good can be consumed, invested in the production of the capital good or stored as inventory. The storage technology transforms one unit of the output good into 1+r units of output good and defines a constant safe rate of return for this economy. The capital good can only be used in the production of the output good and depreciates fully after one period. Formally, the production function takes the general form:

$$y_t = \theta_t f(k_t)$$

where θ_t is a random aggregate productivity shock; it is i.i.d. over time and has mean θ. k_t is the level of capital per head. The production of the output good exhibits constant returns to scale and is subject to an aggregate productivity shock which is not known when the level of the capital stock for the current period is decided.

An exogenous fraction η of agents are entrepreneurs. They have access to the investment technology which is used to produce the capital good. They differ in the efficiency with which they are able to use the investment technology. Formally, entrepreneurs are indexed by an efficiency parameter ω uniformly distributed on the interval [0,1], ranking entrepreneurs from the most efficient to the least efficient. This parameter determines the level of input (in terms of the output good) $x(\omega)$ with $x_\omega > 0$ which is needed to execute one investment project. It says that it is more costly for a less efficient entrepreneur to produce the same amount of capital good.

Entrepreneurs do not consume in the first period so that their savings in the first period are equal to their labour income, $S_t^e = w_t L^e$. Thus, savings (net worth) are procyclical. It is assumed that entrepreneurs' savings are not sufficient to finance the input to the investment project so that entrepreneurs need to borrow from households in order to invest. The return to investment depends on the price (relative to the price of the output good) at which entrepreneurs are able to sell the capital good in the next period, q_{t+1}. They default when the total return to the investment project is lower than the borrowing repayment.

The investment production function is linear in the only input employed in the process which consists simply in the output good. Moreover, the outcome of the investment process is stochastic at the individual level but not at the aggregate level so that there is no uncertainty about the total quantity of capital produced. The distribution of outcomes is identical ex-ante across projects and outcomes are mutually independent. The range of possible outcomes κ_i with their associated probabilities π_i as well as the expected outcome $E(\kappa_i) = \kappa$, are known in advance by all participants.

Households have the choice between investing their savings in the storage technology with constant rate of return r or lending to entrepreneurs. Because they are risk-neutral they only care about the expected return to the investment project. In equilibrium, lenders must be indifferent between lending to entrepreneurs or using the storage technology, therefore their expected return from lending must be equated to the return offered by the storage technology. Lenders must have the insurance that the distribution of outcomes is identical ex-ante and ex-post for borrowing to take place.

2.2.2 The optimal financial contract: illustration with the two-state case

As long as the return to investment is higher than the return to the storage technology, both agents will have an interest in entering in a mutual agreement to finance investment and share the proceeds. The costly state verification model is based on a problem of asymmetric information. Only entrepreneurs can costlessly observe the outcome of their specific project. Although they know the distribution of the projects' outcomes, lenders can discover speific outcomes by activating the auditing technology which absorbs λ units of capital goods. Because entrepreneurs would gain from falsely declaring default, lenders use the threat of random monitoring. have imperfect information about the outcome of entrepreneurs' projects.

The optimal financing contract solves this principal-agent problem. Entrepreneurs' savings S_t^e, the expected future price of capital, q_{t+1} and the return to the storage technology r are treated as parameters. In addition to the amount borrowed and the rate charged by lenders, the solution determines the probability of monitoring by lenders in case of default such that the entrepreneurs' return is maximized, the lenders' constraint is satisfied and entrepreneurs do not lie about the realized outcome of their project.

Here are the main features of the model: the difference between returns to inside and outside funds, the distinction between full and incomplete collateralization and the dependence of the supply of capital on net worth. There are two possible outcomes to the project. State 1 (bad state) occurs with probability π_1 and yields κ_1 units of the capital good. State 2 (good state) occurs with probability π_2 and yields κ_2 units of the capital good (with $\kappa_1 < \kappa_2$). Accordingly, the expected outcome is $\kappa = \pi_1\kappa_1 + \pi_2\kappa_2$. The entrepreneurs consumption payoff is c_1 or c_2 when she announces state 1 or state 2 and is not audited; c_a when she announces state 1, is audited and found not lying; 0 when she announces state 1, is audited and found lying. Let p be the probability of an audit in the bad state and λ be the audit cost.

The optimal contract is characterized by the vector (p, c_1, c_2, c_a) that solves the entrepreneurs' problem. Entrepreneurs choose the values of p, c_a, c_1 and c_2 maximizing their expected future consumption (or profits) subject to a set of constraints: the outside return constraint says that lenders should at least receive the risk-free return r (always binding), the incentive or truth-telling constraint (binds if $p > 0$), the limited liability constraints and the feasibility constraint for p. Formally the problem is written as follows:

$$\max \pi_1 \left[pc_a + (1-p)c_1 \right] + \pi_2 c_2$$

$$\pi_1 \left[q_{t+1}\kappa_1 - p(c_a + q_{t+1}\lambda) - (1-p)c_1 \right] + \pi_2 \left[q_{t+1}\kappa_2 - c_2 \right] \geq r\left(x - S^e\right)$$

$$c_2 \geq (1-p)\left[q_{t+1}(\kappa_2 - \kappa_1) + c_1 \right]$$

$$c_1 \geq 0$$

$$c_a \geq 0$$

$$0 \leq p \leq 1$$

The solution to this problem differs across entrepreneurs according to whether they can default or not, or equivalently whether they are fully collateralized or not.

Inside and outside funds The full collateralization case presents the first channel through which developments in the real economy have an impact on entrepreneurs' access to credit market, by determining the share of entrepreneurs who are fully collateralized. In this case, entrepreneurs' net worth is sufficiently large that they are able to pay lenders their required return even in the bad state:

$$q_{t+1}\kappa_1 \geq \left(x(\omega) - S^e\right)$$

As a result, there is no agency problem, the lender's payoff is independent of the project outcome and $p = 0$. The expected future consumption of the entrepreneur c_{fc} is given by the outside return constraint:

$$c_{fc} = \pi_1 c_1 + \pi_2 c_2 = q_{t+1}\kappa - r\left(x - S^e\right) \quad and \quad \frac{\partial c_{fc}}{\partial S^e} = r$$

The partial derivative says that an additional unit of savings in the first period increases next period consumption by a rate equal to the safe return. Conversely, it is also possible to define the full collateralization level of entrepreneurial saving $S^*(\omega)$, for any given ω, satisfying:

$$S^*(\omega) = x(\omega) - \frac{q_{t+1}}{r}\kappa \quad and \quad \frac{\partial S^*(\omega)}{\partial q_{t+1}} < 0$$

When $S^e \geq S^*(\omega)$, it does not matter for the entrepreneur whether she uses her own funds or borrow in order to invest.

The incomplete collateralization case provides the second channel through which macroeconomic events interfere with financial conditions. In this case, the incentive and liability constraints bind. The optimal auditing probability is given by the outside return condition and the incentive constraint:

$$p = \frac{r\left(x - S^e\right) - q_{t+1}\kappa_1}{\pi_2 q_{t+1}(\kappa_2 - \kappa_1) - \pi_1 q_{t+1}\lambda} \quad with \quad \frac{p}{S^e} < 0 \quad and \quad \frac{p}{q_{t+1}} < 0$$

Assuming $\pi_2(\kappa_2 - \kappa_1) - \pi_1\lambda > 0$, p is always positive when there is incomplete collateralization and is decreasing in the entrepreneurs' net worth S^e. This is the auditing probability that is just sufficient to make the entrepreneur report honestly when the good state occurs. Expected auditing costs $p\pi_1 q_{t+1}\lambda$ determine total agency costs (in terms of the output good) in the economy or the deadweight loss due to the asymmetric information problem. Thus total agency costs are decreasing in the entrepreneurs' net worth S^e. When S^e is low, lenders require a larger share of the expected total return, which is only possible through a reduction in c_2, the second period consumption when the good state occurred; a low c_2 means entrepreneurs have less at risk if they falsely claims the bad outcome, therefore they must be audited more frequently.

Entrepreneurs' expected future consumption c_{ic} is equal to:

$$c_{ic} = \pi_2 c_2 = \alpha \left[q_{t+1}\kappa - r\left(x - S^e\right) - \pi_1 q_{t+1}\lambda\right] \quad with \quad \alpha = \frac{\pi_2(\kappa_2 - \kappa_1)}{\pi_2(\kappa_2 - \kappa_1) - \pi_1\lambda} > 1$$

$$\frac{\partial c_{ic}}{\partial S^e} > r$$

Therefore in the incomplete collateralization case, the return to inside funds S^e (αr) exceeds the return to outside funds (r). Following a marginal increase in net worth S^e, the marginal increase in consumption in the second period is greater for entrepreneurs than for households. The strength of this effect depends on the factors affecting the parameter α, namely the distribution of the idiosyncratic shock and monitoring costs. Because entrepreneurs may belong to either the full collateralization or the incomplete collateralization case, they may or may not have to pledge a greater share of their profits to the lenders. As a result, the average cost of investing in the economy depends on the mixture of these two costs of outside finance (formally it is a weighted average of αr and r).

Thus, the presence of agency costs means that for some entrepreneurs outside funds are more costly than inside funds. In addition whether or not entrepreneurs belong to one or the other category depends on a state variable, S_t^e and the expected future relative price of capital q_{t+1}.

2.2.3 The determination of capital formation

Capital formation in period t is determined jointly by the demand and supply schedules for capital. The first is determined as follows. The expected relative price of capital should equal its marginal return in equilibrium:

$$q_{t+1} = \theta f'(k_{t+1})$$

Given the assumption of CRS in the production of the output good, the demand for capital is downward-sloping in the (q_{t+1}, k_{t+1}) space.

We next derive the schedule for next period's capital stock. The evolution of the capital stock is defined by the following expression:

$$k_{t+1} = (\kappa - h_t \lambda) i_t$$

where i_t is the number of investment projects undertaken in period t and h_t is the fraction of projects initiated in t that are audited. It says that next periods capital stock depends both on the number of investment projects that are initiated and on total agency costs incurred in the process of funding them. We saw that S_t^e and q_{t+1} determined total agency costs. Next we show how q_{t+1} also affects the aggregate level of investment (i_t) or the decision to invest by entrepreneurs.

Entrepreneurs' investment decision Since investment takes the form of discrete and indivisible projects of which each entrepreneur is allotted one, i_t is directly related to the fraction of entrepreneurs who decide to invest. This fraction is lower in the imperfect than in the perfect information case. To show this we first distinguish three types of entrepreneurs: for any period t, let $\overline{\omega}$ and $\overline{\overline{\omega}}$ be the levels of entrepreneurial efficiency that satisfy:

$$q_{t+1}\kappa - rx\left(\overline{\omega}\right) - q_{t+1}\pi_1\lambda = 0$$

$$q_{t+1}\kappa - rx\left(\overline{\overline{\omega}}\right) = 0$$

When $\omega \leq \overline{\omega}$ (good entrepreneurs), expected net return is positive even if the announcement that the bad state has occurred precipitates auditing with probability 1. When $\overline{\omega} < \omega \leq \overline{\overline{\omega}}$ (fair entrepreneurs), expected net returns are positive only if there is no auditing. When $\omega > \overline{\overline{\omega}}$ (poor entrepreneurs), expected net returns are negative even if agency costs are zero. Both $\overline{\omega}$ and $\overline{\overline{\omega}}$ are increasing in q_{t+1}. Given that entrepreneurs will only invest in the project if their expected return is superior to the opportunity costs of funds (rS^e), their optimal choices are summarized in Table 2.2. These choices depend on the full collateralization savings level already defined, S^*, and the level of saving $S^{'}$ at which the return to investment is equal to the return to storage. Poor entrepreneurs always save rather than invest (or become lenders) and good entrepreneurs always choose to invest (and some of them pay agency costs). In terms of the number of investment projects actually realized, the difference with the perfect information case occurs at the level of fair entrepreneurs. Those with the lowest level of saving will not choose to invest (bottom row of the third column) whereas all fair entrepreneurs invest in the benchmark cas.

<table>
<tr><th></th><th>Good ($\omega \leq \overline{\omega}$)</th><th>Fair ($\overline{\omega} < \omega \leq \overline{\overline{\omega}}$)</th><th>Poor ($\omega > \overline{\overline{\omega}}$)</th></tr>
<tr><td>$S^e \geq S^*$</td><td>c_{fc}</td><td>c_{fc}</td><td rowspan="3">No</td></tr>
<tr><td>$S^{'} \leq S^e < S^*$</td><td rowspan="2">c_{ic}</td><td>c_{ic}</td></tr>
<tr><td>$S^e \leq S^{'}$</td><td>No</td></tr>
</table>

Table 2.2: Entrepreneurs' investment decision

To simplify the problem and without loss of generality, the authors introduce the following device to describe the decision of fair entrepreneurs. They first note that their opportunity set is convex between 0 and $S^*(\omega)$, so that risk-neutral entrepreneurs would be happy to enter a fair lottery. The incentive for extra risk-taking here comes from the fact that agency costs are decreasing in the wealth of the agent so that there are increasing returns to wealth over a range. Thus a lottery which pays the full collateralization level of savings S^* with probability $g(\omega) = \frac{S^e}{S^*(\omega)}$ and zero otherwise, improves ex-ante, the fair entrepreneurs expected utility. This simplifying assumption means that fair entrepreneurs become self-financing when they win the lottery and invest.

To summarize, all good entrepreneurs invest but with positive agency costs for some of them and only a fractionof fair entrepreneurs invest without incurring agency costs.

The capital supply schedule The total per capita supply of capital for a given expected relative price is made up of the sum of good and fair entrepreneurs' production:

$$k_{t+1} = \left[\kappa\overline{\omega} - \pi_1\lambda \int_0^{\overline{\omega}} p(\omega)d\omega + \kappa \int_{\overline{\omega}}^{\overline{\overline{\omega}}} g(\omega)d\omega\right]\eta$$

The first two terms of this expression represents capital formation by good entrepreneurs and the final term capital formation by fair entrepreneurs. The probability of good entrepreneurs being audited in the bad state is given by:

$$p(\omega) = \max\left\{\frac{r\left(x - S^e\right) - q_{t+1}\kappa_1}{q_{t+1}\left[\pi_2(\kappa_2 - \kappa_1) - \pi_1\lambda\right]}, 0\right\}$$

with $p(\omega) = 0$ when $S^e \geq S^*(\omega)$. The fraction of fair entrepreneurs who can invest is given by:

$$g(\omega) = \left\{\frac{S_e}{S^*(\omega)}, 1\right\} = \left\{\frac{rS_e}{rx(\omega) - q_{t+1}\kappa_1}, 1\right\}$$

with $g(\omega) = 1$ when $S^e \geq S^*(\omega)$. $g(\omega)$ is increasing in q_{t+1} and S^e. Rewriting the expression for the total supply of capital in terms of its perfect information level:

$$k_{t+1} = \left[\kappa\overline{\overline{\omega}} - \pi_1\lambda \int_0^{\overline{\omega}} p(\omega)d\omega - \kappa \int_{\overline{\omega}}^{\overline{\overline{\omega}}} \left(1 - g(\omega)\right)d\omega\right]\eta$$

The first term represents capital supply in the perfect information case. Several remarks can be made about the difference between the perfect and imperfect information case. First, for a given value of q_{t+1}, entrepreneurs as a whole will supply less capital than in the perfect information case. Second, the capital supply curve is upward-sloping in the (q_{t+1}, k_{t+1}) space and is steeper than in the perfect information case:

$$\frac{\partial k_{t+1}\left[p\left(\omega, S^e, q_{t+1}\right), g\left(\omega, S^e, q_{t+1}\right), \overline{\omega}(q_{t+1}), \overline{\overline{\omega}}(q_{t+1})\right]}{\partial q_{t+1}} > 0$$

As q_{t+1} increases, entrepreneurs' anticipated return rises and a larger fraction of them are fully collateralized. As q_{t+1} becomes big enough, both good and fair entrepreneurs tend to become fully collateralized, $p(\omega)$ and $(1 - g(\omega))$ approach 0 and capital supply approaches its level in the perfect information case. Finally, the capital supply now depends on a period t state variable, S^e: high values of S^e makes capital supply move closer to the perfect information case and are associated with a higher capital supply. Agency costs therefore introduce a channel of dependence of investment on net worth as long as the incentive constraint binds for some entrepreneurs or S^e is not too high.

2.2.4 Persistence and the financial accelerator effect

Agency costs add an additional element of persistence into the model by generating a link between next period capital stock k_{t+1} and the current period entrepreneurial net worth S^e_t, which in turn depends on both the current period capital stock k_t and the productivity shock θ_t. A temporary shock to productivity in the current period, by raising entrepreneurs savings and reducing agency costs for a given level of capital stock, raises investment in the current period and therefore next periods capital stock. By comparison, in the benchmark model with perfect information, a productivity shock has no effect on the level of the capital stock and there is no investment in the current period. In the imperfect information case, the additional increase in activity is therefore generated by a financial accelerator effect. In addition, this mechanism leads to higher savings in the next period, thus potentially giving rise to significant persistent effects. Through the distinction between full and incomplete collateralization, this model is not inconsistent with an asymmetric response to a productivity shock. To explain this intuition we make the following extreme assumptions. We assume the economy starts at the perfect information equilibrium ($k_t = k_{max}$). We further define the minimum value of the productivity shock θ^* generating the level of saving S^e that makes both good and fair entrepreneurs fully collateralized, so that at this level agency costs disappear for these entrepreneurs:

$$\theta^* = \frac{x\left(\overline{\omega}\right) - \left[\frac{q_{t+1}(k_{max})}{r}\right]\kappa_1}{L^e\left[f\left(k_{max}\right) - f'\left[f\left(k_{max}\right)k_{max}\right]\right]}$$

Then, this economy will not react to a realization of the productivity shock above θ^* while generating a persistent fall in economic activity via the financial accelerator mechanism in the opposite case.

One advantage of this model is its ability to give a theoretical foundation to the debt-deflation mechanism described by Fisher (1933). The latter described a situation in which a combination of unindexed debt contracts and unexpected deflation redistributes wealth from debtors to creditors. In the present model, this idea is given shape by assuming an exogenous shock to the distribution of labour endowment between entrepreneurs and lenders while keeping total labour unchanged. Such a negative shock to entrepreneurial labour endowment lowers S^e, raises agency costs and leads to a fall in the capital supply and a higher relative price of capital. Since q_{t+1} rises, debt-deflation cannot explain a stock market crash without introducing additional factors such as aggregate demand externalities or adjustment costs. This model thus proposes the embryo of a framework through which exogenous shocks to financial conditions may initiate or a cyclical fluctuation.

One strength of the general equilibrium approach employed by Bernanke and Gertler (1989) is its greater flexibility. In particular, it is possible to model explicitly the heterogeneity of entrepreneurs, giving rise to nonlinearities in the dynamics of the model. Second, it provides some theoretical foundation for the notion that less efficient entrepreneurs are more sensitive to changes in financial conditions. This hypothesis has been the object of a number of empirical studies investigating the effect of financial frictions for the transmission mechanism of monetary policy (Bernanke and Gertler, 1995). It can also be linked to the more recent work postulating the procyclicality of the risk-taking capacity of the banking sector (Adrian and Shin, 2010; Borio and Zhu, 2012).

Although the financial accelerator and associated persistent effects generated by the model are qualitatively appealing, this framework has nothing to say about their relative strength and magnitude. This is the task Carlstrom and Fuerst (1997) and Bernanke, Gertler and Gilchrist (1999) propose to accomplish. This is achieved at the cost of some simplification.

First, both models adapt the costly state verification approach without stochastic monitoring to a framework that can be calibrated and simulated. In particular, in both cases, the modelling strategy makes clear that the financial friction arises out of the necessity to deal with firms' idiosyncratic risk and not the economy's aggregate risk. Second, for the purpose of aggregation, they both rule out the full collateralization case so that entrepreneurs always need external finance. This is achieved by assuming entrepreneurs' net worth to be sufficiently small, either because they face finite horizons (as in BGG) or because they discount the future more heavily than households (as in CF), so that in both cases, they cannot accumulate net worth indefinitely. Third, for these entrepreneurs, the internal rate of return from running their business is higher than the opportunity cost of the funds invested and they face a premium on external finance. Fourth, entrepreneurs are now identical, therefore ruling out the possibility of endogenous changes in the quality of borrowers. Finally, to facilitate aggregation, both models make two linearity assumptions regarding entrepreneurs' production technologies and lenders' monitoring technologies.

2.3 Carlstrom and Fuerst (1997)

2.3.1 The setup

The contracting problem engages risk-neutral entrepreneurs, who constitute a fraction η of the population, and Capital Mutual Funds (CMFs) through which the savings of risk-averse households are channelled. The contract is drawn in each period after the realization of the aggregate shock and its outcome is known by the end of the same period, therefore eliminating aggregate risk from the problem. Since CMFs are able to diversify entrepreneurs' idiosyncratic risk, they behave effectively as if they were risk-neutral. Entrepreneurial net worth and the aggregate price of capital, defined in the general equilibrium problem, are taken as given in the contracting problem.

Identical entrepreneurs are the producers of the capital good. Entrepreneurs use i units of consumption good as input and produce ωi units of capital good, where ω is a random variable, iid across time and entrepreneurs, with cdf Φ, pdf ϕ and mean $E(\omega) = 1$ and such that $\omega \in [0, \infty)$. ω is only observed by entrepreneurs who have an incentive to lie the true outcome of their investment project. The expected capital output i is a continuous variable, so that all entrepreneurs produce capital goods but differ in the amount they produce. This is guaranteed by the initial constraint that the expected payoff to investment is higher than net worth which is just equal to the opportunity cost of investment for entrepreneurs (participation constraint). CMFs require that the return to lending is no smaller than their opportunity cost of funds, which is just equal to the amount borrowed (zero profit condition).

The principal-agent problem is considerably simplified by the additional assumption that the probability of being monitored in case of default is 1 (no stochastic monitoring) so that there is no informational asymmetry in equilibrium. As a result monitoring costs can be interpreted as default or bankruptcy costs. However, the objective of the optimal contract remains to offer a credible commitment device for the entrepreneur not to lie about the true realization of ω. μi units of capital are used up in the process of monitoring by CMFs. Entrepreneurs finance capital production by using their own savings or net worth n and by borrowing $(i - n)$ consumption goods with repayment $(1 + r^k)(i - n)$ capital goods, where r^k is the lending rate. Therefore the cutoff value of the idiosyncratic shock ω below which they default is given by:

$$\overline{\omega} = \frac{(1 + r^k)(i - n)}{i}$$

When they default, entrepreneurs receive nothing (limited liability constraint) whereas CMFs receive entrepreneurs' profits minus agency costs. Entrepreneurs' expected payoff from capital-good production is equal to:

$$qi \left[\int_{\overline{\omega}}^{\infty} \omega \, \mathrm{d}\Phi(\omega) - (1 - \Phi(\omega)) \, \overline{\omega} \right] = qif(\overline{\omega})$$

$f(\overline{\omega})$ is the expected share of the gross return to entrepreneurs' activity. It is equal to expected profits minus expected repayment costs. CMFs' expected return to lending is equal to:

$$qi \left[\int_{0}^{\overline{\omega}} \omega \, \mathrm{d}\Phi(\omega) - \Phi(\overline{\omega})\mu + (1 - \Phi(\omega)) \, \overline{\omega} \right] = qig(\overline{\omega})$$

$g(\overline{\omega})$ is the expected share of the gross return that goes to the CMFs. It is equal to the sum of the expected return from lending to the entrepreneurs who default (first two terms) plus the expected return from lending to the entrepreneurs who do not default (last term). The expected deadweight cost from producing capital goods caused by the agency problem is equal to:

$$\Phi(\overline{\omega})\mu = 1 - f(\overline{\omega}) - g(\overline{\omega})$$

2.3.2 The optimal contract

Formally, the optimal contract is defined by the pair $(i, \overline{\omega})$ that maximizes entrepreneurs expected gross return:

$$\max_{i,\overline{\omega}} qif(\overline{\omega})$$

subject to lenders' zero-profit condition:

$$qig(\overline{\omega}) \geq i - n$$

and entrepreneurs' participation constraint:

$$qif(\overline{\omega}) \geq n$$

The lenders zero profit-condition allows us to define the expected return to internal funds, which will play a role in the general equilibrium, in the following way:

$$\frac{qif(\overline{\omega})}{n} = \frac{qf(\overline{\omega})}{1 - qg(\overline{\omega})} \geq 1$$

By assumption (entrepreneurs' participation constraint), it is superior or equal to one, the opportunity cost of producing capital goods.

The first-order conditions allow us to derive the following two optimal conditions:

$$q\left[1 - \Phi(\overline{\omega})\mu + \phi\mu\frac{f(\overline{\omega})}{f'(\overline{\omega})}\right] = 1 \quad \rightarrow \quad \overline{\omega}(q) \quad \text{with} \quad \overline{\omega}'(q) > 0$$

$$i = \frac{n}{1 - qg(\overline{\omega})}$$

The first expression defines $\overline{\omega}$ uniquely as an increasing function of q, the aggregate price of capital. As entrepreneurs' activity becomes more profitable, they produce more, therefore increasing agency costs for a given level of net worth, which they are able to do only by accepting a higher default rate given by $\Phi(\overline{\omega})$ (an increasing function of $\overline{\omega}$). The second condition is just the CMFs' zero-profit condition.

2.3.3 Capital production

Substituting the first into the second, the two optimal conditions define the expected quantity of capital produced by each firm $i(q, n)$ which is linear in entrepreneurs' net worth. Aggregating over all entrepreneurs and taking into account monitoring costs, we then have an aggregate supply schedule for new capital goods:

$$I^s(q,n) = i(q,n)\,(1-\Phi(\overline{\omega})\mu) \quad \text{with} \quad I^s_q(q,n) > 0 \quad \text{and} \quad I^s_n(q,n) > 0$$

The aggregate supply of capital goods is increasing in both q and n (aggregate net worth).

Entrepreneurs, similarly to households, rent their capital assets accumulated in the previous period z_t at the rate r_t and sell their labour inelastically for the real wage x_t to the firms producing consumption goods. After consumption goods are produced, entrepreneurs sell their remaining capital assets (net of depreciation) at the price q_t to CMFs. These payments determine their net worth, which is equal to their real wage (entrepreneurs' labour is normalized to one) and the revenue from holding capital assets:

$$n_t = x_t + z_t[q_t(1-\delta) + r_t]$$

δ is the depreciation rate. This expression determines entrepreneurs' supply of capital goods through $i(q,n)$. After capital goods are produced, entrepreneurs who do not default repay their loans to CMFs and decide how much to consume and how much to save, or in other words how much capital assets to retain. Entrepreneurs demand for capital assets is obtained by solving an intertemporal substitution problem, thus determining next period capital asset holdings.

Formally, entrepreneurs maximize their lifetime utility $E_0 \sum_{t=0}^{\infty}(\beta\gamma)^t c^e_t$ subject to the budget constraint $c^e_t + q_t z_{t+1} = qif(\overline{\omega})$. Substituting the budget constraint into the objective function, the problem can be written as:

$$\max_{z_{t+1}} E_o \sum_{t=0}^{\infty} (\beta\gamma)^t \left\{ [x_t + z_t[q_t(1-\delta)+r_t]] \left(\frac{q_t f(\overline{\omega}_t)}{1-q_t g(\overline{\omega}_t)} \right) - q_t z_{t+1} \right\}$$

The first-order condition yields the following Euler equation:

$$q_t = E_t \beta\gamma [q_{t+1}(1-\delta) + r_{t+1}] \left(\frac{q_{t+1} f(\overline{\omega}_{t+1})}{1 - q_{t+1} g(\overline{\omega}_{t+1})} \right)$$

β is households discount rate and γ is the entrepreneurs additional discount rate such that $\gamma \in (0,1)$ and in the steady state, $(\gamma q f(\overline{\omega}))/(1 - qg(\overline{\omega})) = 1$, that is entrepreneurs and households have the same discount factor and they will not postpone consumption indefinitely. This expression says that both entrepreneurs' end-of-period demand for capital assets and the dynamics of the price of capital assets (jointly determined with households' Euler equation) are independent of net worth. Together with the optimal contracting conditions, this implies that an exogenous increase in entrepreneurs net worth reduces agency costs and increases the quantity of capital good produced for any given level of the price of capital.

Because capital does not depreciate fully in each period, the aggregate stock of capital is a state variable in this model This is obtained by aggregating over all entrepreneurs budget constraints:

$$Z_{t+1} = \eta X_t + Z_t[q_t(1-\delta) + r_t]\left(\frac{q_t f(\overline{\omega}_t)}{1 - q_t g(\overline{\omega}_t)}\right) - \eta\frac{c_t^e}{q_t}$$

At the cost of some simplification, it says that entrepreneurs' aggregate capital asset holdings, in terms of units of capital good, are equal to the previous period total profits minus consumption. In turn, this expression determines the next period aggregate net worth.

2.4 Bernanke, Gertler and Gilchrist (1999)

The model postulates the existence of a wedge between the expected aggregate return to capital and the risk-free rate which is referred to as the external finance premium. This key element plays the key pivotal role between the financial structure at the firm level and the real economy.

2.4.1 The setup

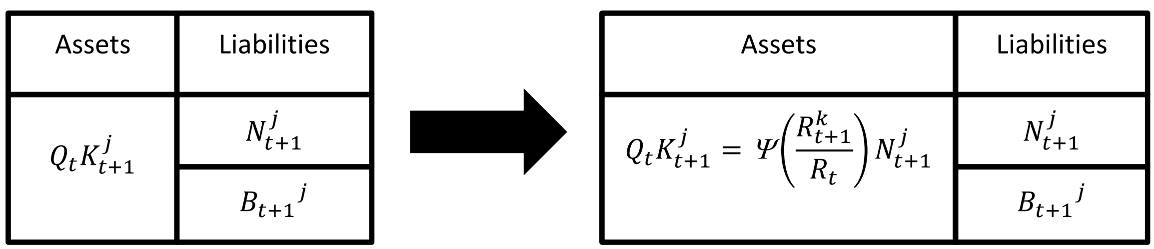


Figure 2-1: Entrepreneur j's balance sheet at the end of period t

Entrepreneurs sell and repurchase their entire stock of capital at the end of each period. They can finance this capital expenditure by either using their own net worth and by borrowing from lenders. Entrepreneurs use the same production technology and have equal access to financial markets but differ in their level of net worth. The ex-post return on the projects depends on two forms of risk: an aggregate risk that affects all entrepreneurs equally and an idiosyncratic risk, which is specific to each entrepreneur. Entrepreneurs gross return to one unit of capital expenditure is given by:

$$\omega R_{t+1}^k Q_t K_{t+1}$$

R_{t+1}^k is the expected gross return to capital and is affected by aggregate risk. $\omega \in [0, \infty)$ is the entrepreneur's idiosyncratic shock with expected value $E(\omega) = 1$[1]. $Q_t K_{t+1}$ denotes the value of the capital stock in real terms, where Q_t is the relative price of capital goods and K_{t+1} is the quantity of capital good purchased. The entrepreneurs' production function exhibits constant

[1] In addition, the distribution of the idiosyncratic shock is assumed to satisfy the following condition: $\frac{\partial(\omega h(\omega))}{\partial \omega} > 0$, where $h(\omega) = \frac{f(\omega)}{1-F(\omega)}$. This condition is satisfied by the lognormal distribution.

returns to scale. The ex-post gross return to capital is observable only at the end of period $t+1$ while entrepreneurs and lenders draw the contract at the end of period t and conditional on the expected gross return to capital.

Entrepreneurs' balance sheets provide a simple way of illustrating the nature of the financial friction. Figure 2-1 is a simple representation of entrepreneur j's balance sheet at the end of period t. N^j_{t+1} is entrepreneur j' s net worth at the end of period t, $B^j_{t+1} = Q_t K^j_{t+1} - N^j_{t+1}$ is the amount borrowed to finance capital expenditure and R^k_{t+1}/R_t is the external finance premium to be defined later. The problem of entrepreneurs at the end of period t is to choose the optimal level of capital expenditure $Q_t K^j_{t+1}$ taking into account the constraint on borrowing imposed by the financial friction. The optimal contract establishes explicitly the relationship between the left-hand side and the right-hand side of the balance sheet of entrepreneurs.

Lenders obtain their funds from risk-averse households who receive with certainty the riskless rate of return on their deposits, R_t. They have the ability to diversify idiosyncratic risk ω perfectly at no cost. In addition, lenders are assumed to be risk-neutral and perfectly competitive. Taken together these assumptions define the constraint on lenders' activity or the lenders' zero-profit condition. This condition determines the lending capacity of financial intermediaries. Formally, it says that the expected return from lending to entrepreneurs should not be inferior to the opportunity cost of the funds borrowed.

The lenders constraint binds for any ex-post realization of aggregate return, in other words there are as many constraints as there are different realizations of the aggregate risk. Since aggregate risk is entirely borne by entrepreneurs lenders never default and ther is no aggregate risk premium.

2.4.2 The optimal contract

Default costs represent a proportion μ of entrepreneurs' payoff that must be paid by lenders in case of default. Default depends on the particular realization of the idiosyncratic shock. The cutoff value of the idiosyncratic shock ($\overline{\omega}$) below which entrepreneurs default is defined as the value of the idiosyncratic shock such that the total gross return to capital is equal to the interest payments to lenders:

$$\overline{\omega}_{t+1} R^k_{t+1} Q_t K_{t+1} = Z_t B_{t+1}$$

Z_t is the contractual gross interest rate. Both $\overline{\omega}_{t+1}$ and Z_t are state-contingent and depend on the actual realization of the aggregate risk. $\overline{\omega}_{t+1}$ is explicitly solved for as the outcome of the contracting problem. A direct and more intuitive counterpart to this variable is the contractual entrepreneurial default rate $F(\overline{\omega}_{t+1})$. Total default costs are determined endogenously through the choice of the default rate.

The financial contract determines the expected shares of the gross return to capital that go to entrepreneurs and borrowers which summarize the respective objectives of both agents.

When $\omega \geq \overline{\omega}^2$, entrepreneurs' profit is equal to the total gross return to capital minus the interest paid to lenders. Their expected share of the gross return to capital is therefore equal to:

$$1 - \Gamma(\overline{\omega}) = \int_{\overline{\omega}}^{\infty} (\omega - \overline{\omega})\, f(\omega)\, \mathrm{d}\omega$$

When $\omega < \overline{\omega}$, entrepreneurs' payoff is zero and lenders get the total gross return to capital minus defaults costs. The latters expected share of the gross return to capital is equal to:

$$\Gamma(\overline{\omega}) - \mu G(\overline{\omega}) = \overline{\omega} \int_{\overline{\omega}}^{\infty} f(\omega)\, \mathrm{d}\omega + \int_{0}^{\overline{\omega}} \omega f(\omega)\, \mathrm{d}\omega - \mu \int_{0}^{\overline{\omega}} \omega f(\omega)\, \mathrm{d}\omega$$

where $\mu G(\overline{\omega}) = \mu \int_0^{\overline{\omega}} \omega f(\omega)\, \mathrm{d}\omega$ is the average share of the realized payoff spent in default costs in the economy.

The lenders zero profit condition can then be written as:

$$[\Gamma(\overline{\omega}) - \mu G(\overline{\omega})]\, R_{t+1}^k Q_t K_{t+1} \geq R_t B_{t+1}$$

For reasonable values of R_{t+1}^k and $\overline{\omega}$, this condition will hold with equality. It says that lenders equate the gross opportunity cost of funding (applying at the margin) to the gross return to lending (which is an average concept).

Entrepreneurs Taking R_{t+1}^k, R_t, Q_t and N_{t+1} and the distribution of the idiosyncratic risk as given, entrepreneurs choose both the quantity of capital K_{t+1} and the cutoff value of the idiosyncratic shock $\overline{\omega}$ that maximize their expected return (profit) subject to the lenders zero-profit condition:

$$\max_{K_{t+1}, \overline{\omega}} E_t(V_{t+1}) = (1 - \Gamma(\overline{\omega})) R_{t+1}^k Q_t K_{t+1}$$

subject to

$$[\Gamma(\overline{\omega}) - \mu G(\overline{\omega})]\, R_{t+1}^k Q_t K_{t+1} = R_t (Q_t K_{t+1} - N_{t+1})$$

This problem can conveniently be rewritten in terms of the external finance premium R_{t+1}^k / R_t and the inverse of entrepreneurs' leverage ratio $k_{t+1} = (Q_t K_{t+1})/N_{t+1}$. The external finance premium is the ratio of the expected return to capital (inside funds) to the risk-free rate (outside funds). The problem is reformulated as:

$$\max_{k_{t+1}, \overline{\omega}} (1 - \Gamma(\overline{\omega})) \frac{R_{t+1}^k}{R_t} k_{t+1}$$

[2] To simplify the exposition, we will no longer indicate the time subscript for this variable.

subject to

$$[\Gamma(\overline{\omega}) - \mu G(\overline{\omega})] \frac{R^k_{t+1}}{R_t} k_{t+1} = k_{t+1} - 1$$

Thus the two choice variables, the default rate and entrepreneurs leverage, are not a function of net worth and only depend on the external finance premium. The contract is defined as the optimal pair $(k_{t+1}, \overline{\omega})$ solving this problem.

First, we solve for the optimal value of k^*_{t+1} as a function of the expected gross return to capital and the associated expected cutoff value $\overline{\omega}$. Second, conditional on k^*_{t+1}, a schedule for the state-contingent value of the optimal cutoff value $\overline{\omega}^*$ is determined by the lenders constraint.

Solving the problem by substitution yields two first-order conditions[3]:

$$\frac{1 - F(\overline{\omega})}{1 - \Gamma(\overline{\omega})} = \frac{1 - F(\overline{\omega}) - \mu F'(\overline{\omega})}{1 - \frac{R^k_{t+1}}{R_t}[\Gamma(\overline{\omega}) - \mu G(\overline{\omega})]}$$

$$k_{t+1} = \frac{1}{1 - \frac{R^k_{t+1}}{R_t}[\Gamma(\overline{\omega}) - \mu G(\overline{\omega})]}$$

BGG (1999) showed that, provided the external finance premium is superior to one and is not too large, these two conditions guarantee a one-to-one mapping between, on one hand the optimal cutoff value and the external finance premium, and on the other hand the optimal value of the inverse of the leverage ratio and the optimal cutoff value. These two results in turn imply that at the optimum, the following relationship must hold:

$$k^j_{t+1} = \Psi\left(\overline{\omega}\left(\frac{R^k_{t+1}}{R_t}\right)\right) = \Psi\left(\frac{R^k_{t+1}}{R_t}\right) \quad \text{with} \quad \Psi(1) = 1 \quad \text{and} \quad \Psi'\left(\frac{R^k_{t+1}}{R_t}\right) > 0$$

This expression says that in equilibrium the inverse of the leverage ratio of entrepreneurs is uniquely determined by the expected value of the external finance premium.

The external finance schedule(1) Entrepreneurs' optimal capital purchases are therefore subject to the financial constraint:

$$Q_t K^j_{t+1} = k^j_{t+1} N^j_{t+1} = \Psi\left(\frac{R^k_{t+1}}{R_t}\right) N^j_{t+1} \quad \text{with} \quad \Psi'\left(\frac{R^k_{t+1}}{R_t}\right) > 0$$

According to this financial constraint, capital expenditure and net worth are linearly related to each other.

[3]Note that when $\mu = 0$, this first-order condition reduces to $R^k_{t+1} = R_t$.

2.4.3 The determination of aggregate capital expenditures

The external finance schedule (2) At the end of each period, entrepreneurs sell their remaining capital stock to capital-producing firms. The solution to the contracting problem determines the capital supply schedule (or external finance schedule) which jointly determines the next period level of capital stock. In the previous section, we derived the external finance schedule faced by each individual entrepreneur. By the linearity assumption, the external finance schedule for the entrepreneurial sector as a whole is simply: $Q_t K_{t+1} = \Psi\left(\frac{R^k_{t+1}}{R_t}\right) N_{t+1}$. Becasue the contract is determined conditional on the realization of aggregate unicertainty, the external finance schedule can equivalently be rewritten in expectational terms as:

$$E(R^k_{t+1}) = \chi\left(\frac{N_{t+1}}{Q_t K_{t+1}}\right) \quad \text{with} \quad \chi'\left(\frac{N_{t+1}}{Q_t K_{t+1}}\right) < 0$$

This expression establishes the critical link between the microeconomic problem and the general equilibrium model where R^k_{t+1} is determined endogenously. This external finance schedule determines the lowest value of the expected return to capital necessary to purchase a given amount of capital expenditure ($Q_t K_{t+1}$). It is inversely proportional to entrepreneurs net worth. The internal rate of return (left-hand side) is determined by the marginal product of capital and capital gains (see Appendix A), allowing changes in asset prices to amplify the effects of the financial friction.

Net worth is a state variable and therefore plays a key role in generating persistence. Understanding its determinants and its interplay with macroeconomic conditions is critical in understanding how this microeconomic problem can have significant macroeconomic consequences. Aggregate entrepreneurial net worth is defined as:

$$N_{t+1} = \lambda V_t + W^e_t$$

W^e_t is the entrepreneurs real wage in period t and V_t is aggregate entrepreneurs profit at the end of period t and is defined by:

$$V_t = R^k_t Q_{t-1} K_{t+1} - \left[R_{t-1} + \frac{\mu G(\overline{\omega}_t) R^k_t Q_{t-1} K_t}{B_t}\right] B_t$$

The authors used a simplified expression for V_t[4]. Because of the financial friction, in this model, the competitive entrepreneurial sector is able to make a profit in equilibrium.

Aggregate net worth is a state variable; its current level depends on its level in the previous

[4] In theory aggregate net worth should be equal the sum of (1) the net worth of those who defaulted (with probability $F(\overline{\omega})$), which is equal to their real wage, (2) the net worth of those who were just born and whose net worth is also equal to their real wage and finally (3) the net worth of those who did not default and whose net worth were described in section ? (in addition these entrepreneurs have different levels of net worth depending on whether they defaulted and/or when they were born).

period and its dynamics are given by the following difference equation:

$$N_{t+1} = \lambda[(1 - \mu G(\overline{\omega}_t))R_t^K Q_{t-1}K_t - R_t(Q_{t-1}K_t - N_t)] + W_t^e$$

The parameter λ imposes an upper bound on the level of net worth. Because entrepreneurs are leveraged, in other words because they fund themselves externally, their profits respond more than one for one to a change in R_t^k. The behaviour of the external finance premium over the business cycle, and equivalently of the relative price of capital, will therefore play a significant role in the dynamics of capital accumulation.

2.5 Kiyotaki and Moore (1997)

KM develop a general equilibrium model in which the presence of endogenous credit constraints explain why exogenous productivity shocks may have persistent effects on the real economy. The microeconomic motivation for the existence of the financial friction, however, leads to credit rationing, thus shifting the focus from interest rate spreads to asset prices. KM give a prominent role to the interaction between credit frictions and the dynamics of asset prices over the business cycles in generating the financial accelerator effect. In particular, KM show how the persistence and amplification effects mutually reinforce each other[5]. This model is also able to give shape to the idea that sector-specific shocks may have real effects, similar to the income redistribution effects in BG.

2.5.1 The basic setup

The focus of the paper is on the role of capital assets as collateral in the contract between borrowers and lenders in an economy with a fixed supply of capital. Aggregate uncertainty is not built into this model explicitly. However, it is assumed that one-time exogenous productivity shocks drive temporarily the system out of the steady state. The anticipation of such shocks also motivates the financial friction.

The producers

Economic activity consists in using the capital asset, land, which is in fixed supply, $\overline{K}$, to produce a nondurable consumption good, fruit (numeraire good) in an environment without aggregate uncertainty and with rational expectations. There are three markets, for land, fruit and credit. Fruits are either consumed, exchanged for land at the price $1/q_t$[6], or for R_t units of fruit in the next period.

There are two categories of infinitely-lived, risk-neutral producers-consumers in the economy, farmers (with size normalized to 1) and gatherers (with size m). In each period, they both produce fruit using their predetermined quantity of land and they both choose how much to

[5] KM make the distinction between the static multiplier process developed by Schleifer and Vishny (1992) and the intertemporal multiplier model they develop.

[6] In other words, one unit of land costs q_t units of fruits.

consume, how much to save and how much to invest in land. They differ in the way they discount future consumption and in their production technologies.

Objective functions

Formally, farmers' and gatherers' respective objective functions are:

$$E_t\left[\sum_{s=0}^{\infty}\beta^s x_{t+s}\right]$$

and

$$E_t\left[\sum_{s=0}^{\infty}\beta'^s x_{t+s}^G\right]$$

where x_{t+s} and x_{t+s}^G are their respective consumption of fruit and β and β', their respective discount factors. It is assumed that $\beta < \beta'$ (assumption 1), so that farmers are relatively more impatient[7].

Production functions

Farmers' production function is linear in their land holdings . A fraction of their output is assumed to be non-tradable so that it must be consumed by farmers. Formally, output in the next period is:

$$y_{t+1} = (a+c)k_t$$

where ak_t is the tradable output, ck_t is the non-tradable output and k_t is farmers land holdings. It is assumed that $a/(a+c) < \beta$ (assumption 2), where the left-hand side is the maximum fraction of their output that can be reinvested or saved by farmers.

Gatherers' production function exhibits decreasing returns to scale:

$$y_{t+1}^G = G(k_t^G) \quad \text{and} \quad G' > 0, \quad G'' < 0 \quad \text{and} \quad G'(0) > aR > G'\left(\frac{\overline{K}}{m}\right)$$

where aR is the gatherers' steady-state marginal product of land. All of their output is tradable.

Budget constraints

Farmers' and gatherers' respective budget constraints are given by:

$$q_t(k_t - k_{t-1}) + Rb_{t-1} = ak_{t-1} + b_t$$

[7]According to KM, this condition ensures that there is no steady-state equilibrium in which the farmers' credit constraints are not binding.

and

$$q_t(k_t^G - k_{t-1}^G) + Rb_{t-1}^G + x_t^G = G(k_{t-1}^G) + b_t^G$$

where q_t is the price of land and b_t and b_t^G are the respective amounts of debt incurred in period t. In equilibrium, farmers' and gatherers' aggregate levels of debt must sum to zero, so that some will be net borrowers and the others net creditors.

The financial friction

Farmers' and gatherers' streams of income in period $t+1$ are equal to the sum their income and the value of their land assets, which is also the liquidation value of land holdings:

$$(a + c + q_{t+1})k_t \quad \text{and} \quad (G'(k_t^G) + q_{t+1})k_t^G$$

Contrary to gatherers, farmers have the possibility to withdraw their labour after the contract has been agreed and before production has started[8]. Assume farmers borrow $[(a + q_{t+1})k_t]/R_t$ to finance their land purchases at the end of period t. In period $t+1$, if farmers do not renege on their contract, they consume ck_t and use the remaining of their income to repay their loan. However, at the beginning of period $t+1$, before production starts, farmers may threaten their creditors to repudiate their contract unless their outstanding debt is reduced by any amount Δa ($\Delta a < a$). Creditors are then facing the following alternative. If farmers withdraw their labour, then the total return from their land acquisition in the next period is just equal to the total value of their land assets $q_{t+1}k_t$ which is the maximum amount they can repay creditors. If creditors agree to reduce farmers outstanding debt by Δa and farmers do not renege on the new contract, loan repayments will be $(a - \Delta a + q_{t+1})k_t > q_{t+1}k_t$. Thus creditors would be willing to renegotiate the contract.

Farmers' creditors therefore incur a form of idiosyncratic risk if borrowing repayments exceed the liquidation value of land. However, it is possible to adapt this framework so that aggregate risk is given a central role. Suppose that, at the beginning of period $t+1$, before production starts, there is a negative productivity shock reducing farmers' tradable output by the amount Δa. This would justify their demand to renegociate their contract.

To eliminate this risk, creditors impose the constraint that borrowing repayments may not exceed the collateral value of land:

$$R_t b_t \leq q_{t+1}k_t$$

Gatherers are not subject to this constraint by assumption. However, for farmers, this amounts to reducing the problem to the full collateralization case, so that this model does not generate a premium on external finance and the interest paid on debt is the standard equilibrium interest rate. However, it introduces a wedge between farmers' inside value of land (as represented by their income stream) and its outside value of land (liquidation value). A final assumption rules

[8]This follows from the notion of the inalienability of farmers' human capital developed in Hart and Moore (1994) and the fact that farmers are assumed to have specific skills.

out bubbles in the price of land (assumption 3):

$$\lim_{s\to\infty} E_t(R^{-s} q_{t+s}) = 0$$

This setup implies that farmers are leveraged while gatherers are not and this plays an important role in explaining the interaction between collateral values and credit constraints. In equilibrium, farmers' marginal product of land is greater than gatherers' marginal product, implying that the equilibrium level of aggregate output is inefficient. In addition, farmers' internal return to capital is greater than the opportunity cost of funds, in other words, farmers' Tobin's q is higher than the market price of capital.

2.5.2 The general equilibrium

The equilibrium in this model is completely defined by the infinite sequence of land prices, allocations of land between farmers and gatherers, as well as their debt and consumption choices.

Gatherers

Gatherers solve a standard lifetime utility maximizing problem subject to a budget constraint. Substituting the latter into the former:

$$\max_{k_{t+s}^G} E_t \left[\sum_{s=0}^{\infty} \beta'^s \left\{ G(k_{t+s-1}^G) + b_{t+s}^G - Rb_{t+s-1}^G - q_{t+s}(k_{t+s}^G - k_{t+s-1}^G) \right\} \right]$$

The first-order condition yields the following optimal condition:

$$\frac{G'(k_t^G)}{R} = q_t - \frac{q_{t+1}}{R} = u_t$$

The equilibrium interest rate is such that gatherers are indifferent between saving a marginal unit of fruit or consuming it. Because gatherers' utility is linear, the interest rate is just equal to $R = 1/\beta'$. The discounted marginal product of capital is equated to the opportunity cost of holding the land. It says that in equilibrium, gatherers are indifferent between investing a marginal unit of land asset into production or selling it at price q_t (to farmers) in period t, therefore earning interest R with the proceed, and buying it back at price q_{t+1} in period $t+1$. This condition determines gatherers' demand for land. Gatherers' aggregate demand for land is $k_t^G m$ is therefore a function of q_t and q_{t+1} only.

Farmers

In the neighbourhood of the steady state, farmers prefer to borrow up to the maximum allowed and invest in land (see appendix for demonstration)[9]. Thus farmers' consumption in each

[9] In addition, assumption 3 guarantees that there is a locally unique perfect foresight equilibrium path in the neighbourhood of the steady state (p222).

period is just equal to their non-tradable fruit output:

$$x_t = ck_{t-1}$$

Their level of debt is given by the borrowing constraint:

$$b_t = \frac{q_{t+1}k_t}{R}$$

Their demand for capital is given by their budget constraint:

$$\left(q_t - \frac{q_{t+1}}{R}\right)k_t = (a + q_t)k_{t-1} - Rb_{t-1}$$

The right-hand side represents farmers net worth in period t. It is equal to the sum of the tradable output and the collateral value of land minus borrowing repayments. The left-hand side the amount of land purchased by farmers that has to be financed by their net worth. $q_t - q_{t+1}/R = u_t$ is the downpayment required to buy one unit of land. This condition says that farmers use all of their net worth to finance their capital purchase.

Farmers' aggregate demand for land and debt is obtained by aggregating individual demands:

$$K_t = \left(\frac{1}{u_t}\right)[(a + q_t)K_{t-1} - RB_{t-1}]$$
$$B_t = \frac{q_{t+1}K_t}{R}$$

Consider a simultaneous one per cent increase in the prices of land in periods t and $t+1$, leading to a one per cent increase in u_t. This shifts farmers' demand for land through three different channels. First, it has a negative impact by raising the price of land in the current period. Second and third, it has a positive impact by increasing the current and future collateral value of land, which respectively boosts net worth and relaxes the credit constraint. Defining net worth as $n_t = (a + q_t)K_{t-1} - RB_{t-1}$, its elasticity with respect to asset prices is given by:

$$\epsilon_q^n = \frac{q_tK_{t-1}}{n_t} = \frac{RB_{t-1}}{aK_{t-1}} > 1 \quad \text{when} \quad RB_{t-1} > aK_{t-1}$$

where the last inequality holds in equilibrium. Thus the nominator in the expression for the aggregate demand for land increases by more than the denominator and farmers aggregate demand for land increases. This result follows from the fact that farmers borrow in order to buy land[10]. Farmers' leverage ratio in each period can be recovered from their optimal demand for land: $(q_tk_t)/n_t = q_t/u_t$. By comparison, from their optimal condition for land holdings, it is easily seen that gatherers react to the same shock by reducing their demand for land assets.

[10]Suppose farmers did not borrow in the previous period so that their net worth is just equal to $(a+q_t)K_{t-1}$, then the elasticity of net worth with respect to asset price would be: $\epsilon_q^n = (q_tK_{t-1})/n_t = (q_tK_{t-1})/[(a + q_t)K_{t-1}] < 1$. Farmers' demand for land would tend to fall.

Market clearing conditions and the price of capital

The land market clearing condition states that the sum of farmers' and gatherers' land holdings must be equal to total land supply: $\overline{K} = K + mk_t^G$. Substituting this condition into the gatherers' optimal condition:

$$\frac{1}{R}G'\left[\frac{1}{m}(\overline{K} - K_t)\right] = u_t(K_t) \quad \text{with} \quad u_t'(K_t) > 0$$

For the land market to clear in each period, an increase (decrease) in farmers' (gatherers') demand for land should be associated with an increase in gatherers' opportunity cost of holding land. The determination of the path for the allocation of land fully determines consumption and debt sequences. In equilibrium, farmers are net borrowers so that gatherers must be net creditors in equilibrium.

The land market condition and the definition of the opportunity cost of land holdings for gatherers imply the following expression for the price of land assets:

$$q_t = \sum_{s=0}^{\infty} R^{-s} u(K_{t+s})$$

The price of land in period t is equal to the present value of the infinite sequence of capital gains generated by the possession of one unit of land in period t, as priced by the market.

The steady state

In the steady state farmers land holdings and debt are constant as well as the price of land, so that the collateral value of land is equal to debt and tradable output must be equal to debt interest payments. The steady state values for u_t and q_t are easily obtained from farmers' aggregate demand for land condition and are respectively equal to $u^* = a$ and $q^* = (aR)/(R-1)$. Gatherers' steady state marginal return to land holding from their optimal condition for land holding is given by: $G'(k^G) = G'[(1/m)(\overline{K} - K^*)] = aR$. Assumptions 1 and 2 imply that farmers' marginal return to capital is greater than gatherers' in the steady state:

$$a + c > aR$$

Thus a shock redistributing land from gatherers to farmers would lead to an increase in aggregate output. Noting that farmers' net worth is equal to aK, this inequality, rewritten as $(a+c)/a > R$, says that farmers' return to internal fund is greater than their opportunity cost of funds. Finally, farmers' steady state leverage ratio is equal to $q^*/u^* = R/(R-1)$.

The steady state value of the price of land can be rewritten as:

$$q = \frac{G'(k^G)}{R-1}$$

Asset price q is equal to the present discounted value of the future stream of revenue generated by gatherers for holding their marginal unit of land in each period. In other words, the steady value of q is equal to gatherers' Tobin's q[11].

The effect of a productivity shock

To study how an exogenous one-time shock to productivity is transmitted in this economy, one can make these simplifying assumptions: the economy starts at the steady state and both farmers' and gatherers' expected production is multiplied by $(1+\Delta)$ in period t, after farmers have made their labour decision[12]. Substituting the land market clearing condition into the expression for farmers' aggregate demand for land in period t and $t+s$ $(s \geq 1)$ yields the following sequence of difference equations:

$$u(K_t)K_t = [a + \Delta a + q_t - q^*]K^*$$
$$u(K_{t+s})K_{t+s} = aK_{t+s-1}$$

The first term in each expression nests the requirement (in terms of their opportunity cost) demanded by gatherers to accommodate any level of farmers land holdings. The second expression is the usual condition. The first expression shows that the shock changes farmers net worth by affecting both the level of their current tradable output and the collateral value of their current land holdings. Linearizing around the steady state:

$$\left(\frac{1}{\eta}+1\right)\widehat{K}_t = \Delta + \left(\frac{R}{R-1}\right)\widehat{q}_t$$
$$\left(\frac{1}{\eta}+1\right)\widehat{K}_{t+s} = \widehat{K}_{t+s-1}$$

The respective changes in farmers' land holdings $\widehat{K}_t$ and land prices $\widehat{q}_t$ are expressed in log-deviations from their steady state values. The parameter $\frac{1}{\eta}$ measures by how much gatherers require the opportunity cost of land to change in order to accommodate a marginal change in farmers' demand for land:

$$\frac{1}{\eta} = \frac{\mathrm{d}\log U(K^*)}{\mathrm{d}\log K^*} = -\frac{\mathrm{d}\log G'(k'^*)}{\mathrm{d}\log k'^*}\frac{K_t}{\overline{K}-K_t} > 0$$

wher $1/\eta$ is the elasticity of the gatherers' marginal product of land with respect to their land holdings times the ratio of the farmers' to the gatherers' landholdings in the steady state. In other words, η is the elasticity of the residual supply of land to the farmers with respect to the user cost evaluated at the steady state[13].

The first linearized equation shows how the double impact of a positive productivity shock

[11] In an environment without credit constraints, farmers' and gatherers' marginal products of capital would be equated and the steady state value of the price of land would be $q = (a+c)/(R-1) > q = G'(k^G)/(R-1)$.

[12] The same process could be applied to study the effect of an exogenous redistribution of wealth between gatherers and farmers, in which case this exogenous increase in land productivity would only be applied to farmers' production.

[13] This parameter, which can be interpreted as a measure of how liquid capital assets are, is therefore not allowed to vary over the credit cycle in the linearized version of the model.

on the level of tradable output (direct effect) and on the collateral value of land holdings (indirect effect) is transmitted to the current change in farmers' land holdings. In particular, the strength of the collateral value effect depends on the steady state value of the leverage ratio. In the absence of borrowing constraint, the price of land would be fixed at the same level over the business cycle and would therefore play no part in the dynamics of real output. Finally, these positive effects are in part crowded out by the implied cost necessary to induce gatherers to give up some of their land assets (left hand side). The solution to this linearized system of equations for the current period change in farmers' land holdings and land price illustrates further these intuitions and gives a sense of the magnitude of the effects implied by the parameters of the model:

$$\widehat{K}_t = \left(\frac{\eta}{1+\eta}\right)\left[\Delta + \left(\frac{R}{R-1}\right)\frac{1}{\eta}\Delta\right] \quad \text{and} \quad \widehat{q}_t = \frac{1}{\eta}\Delta$$

We note that the more elastic the supply of lands to farmers is (or the more liquid are these assets), the less marked is the amplification effect associated with the asset price channel[14].

The second linearized equation shows how a temporary shock in period t still modifies the system in the subsequent periods (persistence effect) by durably affecting farmers' levels of net worth and their required down payments.

2.5.3 Static versus dynamic multiplier

This basic model allows distinguishing between a static and a dynamic multiplier effect associated with the behaviour of asset prices. This difference may be economically significant. The intuition behind this idea is that because agents expect temporary shocks to have persistent effects (by affecting durably farmers' net worth), they also expect them to generate persistent fluctuations in asset prices, which in turn feed back into the current valuation of assets and therefore net worth.

Suppose the land price returns to its steady state value after one period so that $\widehat{K}_{t+s} = 0$ for $s \geq 1$. Then the solution to the linearized system for the current period change in farmers' land holdings and land price is:

$$\widehat{K}'_t = \Delta \quad \text{and} \quad \widehat{q}'_t = \frac{1}{\eta}\left(\frac{R-1}{R}\right)\Delta$$

The additional percentage deviations of land holdings and land price from the steady state due the dynamic multiplier are respectively equal to $(1+\eta)(R-1)$ and $1/(R-1)$ times those due to the static multiplier.

[14]KM note that due to the large multiplier effects, the nonlinear system may have multiple equilibria.

2.6 Conclusion

In this section we make several remarks highlighting the differences between the models of Carlstrom and Fuerst (1997), Bernanke, Gertler and Gilchrist (1999) and Kiyotaki and Moore (1997) and their limitations, in particular in terms of their ability to offer a coherent framework to think about the interaction between borrowers' financial conditions, risk premia, asset prices and bank risk-taking.

Aggregate uncertainty and associated macroeconomic risk premia play little role in these models whose focus is on the external finance premium or the difference between the internal and external returns to funds and to a lesser extent on probability of default. In CF, aggregate risk is eliminated from the debt contract by making it an intra-period contract. The optimal contract yields a unique cutoff value for the idiosyncratic shock ($\overline{\omega}$), or equivalently a unique default rate $\Phi(\overline{\omega})$, from which the external risk premium charged by lenders r^k can be recovered. In BGG, aggregate risk is included ex-post in the outcome of the agency problem. The solution to the contracting problem yields the ex-ante leverage ratio and state-contingent schedule for the default rate $F(\overline{\omega})$, and therefore the premium on external finance $(\mu G(\overline{\omega})R_t^k Q_{t-1}K_t)/B_t$. This risk premium depends ex-post on aggregate uncertainty. The ex-post risk premium only affects the economic rents of entrepreneurs. If entrepreneurs are banks and lenders are the mutual funds that fund them, then this premium can be interpreted as a measure of credit risk. The endogenous variation in the default rate should feed into the real economy through the dynamics of net worth, therefore amplifying further the business cycle fluctuation. Thus, a role is given to the interaction between bankruptcy rates, risk premia and macroeconomic conditions.

There are differences in the role assigned to asset prices in the three models. In CF, a positive shock to economic activity which boosts firms' net worth leads to an increase in the production of capital goods I and a fall in the price of capital q. This has the offsetting effect of depressing net worth, therefore dampening the initial impact of the rise in net worth. Thus in CF, the price of capital is countercyclical. In BGG, an exogenous increase in net worth leads to an increase in equilibrium capital expenditures and therefore aggregate investment, pushing up the price of capital because of the presence of adjustment cost in the production of capital goods. This generates capital gains on current capital asset holdings, increasing current profits and further boosting net worth and capital expenditure (while depressing by a lesser extent the demand schedule for capital because of the anticipation of future capital losses). Thus in BGG the price of capital assets is procyclical.

A related distinction is between the models generating persistence and those generating amplification effects. CF, following Bernanke and Gertler (1989), show how an exogenous shock to net worth can have persistent effects on aggregate dynamics. BGG add to this persistence an amplification effect of asset prices. One of the strength of KM is to develop a more parsimonious model that is also able to generate an amplification effect of asset prices. The main contribution of this model to the literature resides in the role it assigns to asset prices. In KM, capital prices in each period reflect the infinite sequence of future capital gains on asset holdings (for gatherers). A positive productivity shock that increases persistently farmers' demand for land necessitates a persistent rise in gatherers' opportunity cost of capital asset holdings, and therefore asset (land) prices. This, in turn, boosts further the initial rise in asset prices, which relaxes even more farmers' credit constraint on impact. This mechanism gives rise to a dynamic multiplier effect, as opposed to the static multiplier effect present in BGG. Thus, asset prices in KM are procyclical and they amplify the financial accelerator effect. However, they do not reflect fundamentals, but the economic conditions faced by gatherers (not farmers), thus creating the conditions for bubbles to form.

These distinctions may have different implications for the economic significance of financial accelerator mechanisms and their ability to generate nonlinear dynamics.

Appendix A: Bernanke, Gertler and Gilchrist (1999)

Entrepreneur j's net worth

Capital expenditures in each period depend on entrepreneurs' net worth N_{t+1}^j. Entrepreneur j's net worth is defined as the sum of her profits realized in period t and her real wage:

$$N_{t+1}^j = V_t + W_t^e$$

W_t^e is the entrepreneurs real wage in period t. Entrepreneurs' profit at the end of period t, V_t, is the realized return to capital investment for entrepreneurs who do not default and zero otherwise. It is equal to the gross return to capital investment from period $t-1$ to period t less borrowing repayments[15]:

$$V_t^j = \omega^j R_t^k Q_{t-1} K_{t+1}^j - \left[R_{t-1} + \frac{\mu G(\overline{\omega}_t) R_t^k Q_{t-1} K_t}{B_t} \right] B_t^j$$

The term in bracket corresponds to the actual interest the entrepreneurs who do not default are charged by lenders: $Z_{t-1}^j = R_{t-1} + (\mu G(\overline{\omega}_t) R_t^k Q_{t-1} K_t)/B_t$. The second term is the ratio of expected default costs (in the economy) to average quantity borrowed and is the actual premium paid by entrepreneurs for funding their investment expenditure externally. Its actual value, as for the cutoff value of the idiosyncratic shock, depends on the particular realization aggregate risk. Finally, because net worth depends on the amount of debt incurred in the previous period, it depends on the previous period level of net worth.

The determination of the internal return to capital

To understand how the internal rate of return is eventually determined, we make a digression into the general equilibrium problem. The expected return per unit of capital expenditure R_t^k is determined by the marginal product of capital and capital gains and is expressed in terms of the final retail price. The demand curve for capital expenditures is given by:

$$E(R_{t+1}^k) = E\left[\frac{\frac{1}{X_{t+1}} \frac{\alpha Y_{t+1}}{K_{t+1}} + Q_{t+1}(1-\delta)}{Q_t} \right]$$

wher δ is the rate of depreciation of capital, X_t is the retailers markup so that $1/X_t$ is the relative price of wholesale goods and Q_t is the relative price of capital goods (in terms of the numeraire retail good) charged by capital-producing firms. Competitive capital-producing firms buy the final good and rent capital[16] to produce new capital goods using the following

[15] This expression is found by substituting the lenders constraint into the entrepreneurs expected return. It closely resembles the objective function defined to derive the solution to the contracting problem, the difference being that instead of expected profit, it refers to ex-post profit (after the realization of aggregate risk).

[16] More precisely, they buy entrepreneurs entire capital stock at a price $\overline{Q}_t$ and sell it back at the price Q_t; this defines the implicit rental rate as $\overline{Q}_t - Q_t$. However, this difference is treated as second-order in the solution to the problem.

technology:

$$\Phi\left(\frac{I_t}{K_t}\right)K_t \quad \text{with} \quad \Phi'\left(\frac{I_t}{K_t}\right) > 0, \Phi''\left(\frac{I_t}{K_t}\right) < 0 \quad \text{and} \quad \Phi(0) = 0$$

I_t is aggregate investment expenditures. This production function exhibits CRS and increasing marginal adjustment costs: as the ratio of real investment to capital stock increases, the marginal return of the investment technology falls. The price of capital (Tobins q) is given by:

$$Q_t = \frac{1}{\Phi'\left(\frac{I_t}{K_t}\right)} \quad \text{and} \quad \frac{\partial Q_t}{\partial (I_t/K_t)} = -\frac{\Phi'}{\Phi'^2} > 0$$

It is increasing in the ratio of the aggregate level of investment to aggregate capital stock, and therefore real activity. Movements in the relative price of capital lead to exogenous shifts in the schedule for the demand for capital.

Appendix B: Kiyotaki and Moore (1997)

Proof of the proposition on farmers optimal behaviour in the neighbourhood of the steady state: using the principle of unimprovability, we consider three possible uses of a farmers' marginal unit of tradable output in period t. In the first, they invest it in $1/u_t$ units of land. In the second, they save it at the interest R in the first period and then invest the entire proceed in land assets. In the third, they just consume it. Next, we compare the expected utility farmers receive under these three consumption paths, making use of the fact that the steady state value of u_t is given by $u^* = a$.

Consumption path 1 (investment): $0 + \beta\frac{c}{a} + \beta^2\frac{c}{a} + \beta^3\frac{c}{a} + \ldots = \left(\frac{\beta}{1-\beta}\right)\frac{c}{a}$

Consumption path 2 (saving): $0 + 0 + \beta^2 R\frac{c}{a} + \beta^3 R\frac{c}{a} + \ldots = \left(\frac{\beta^2}{1-\beta}\right)\frac{1}{\beta'}\frac{c}{a}$

Consumption path 3 (consumption): 1

Recalling assumptions 1 and 2, $\beta < \beta'$ and $c > [(1-\beta)/\beta]a$, we conclude that the first consumption path, offered by investment, strictly dominates the other two. The investment opportunity described here is one where the marginal unit of tradable output is used to cover the down payment (or the haircut) necessary to purchase land while the other share is borrowed using the future value of this investment as collateral. Thus this result also shows that farmers' credit constraints are always binding.

Chapter 3

Financial accelerator models: banks' balance sheet

3.1 Holmstrom and Tirole (1997)

Holmstrom and Tirole (1997) construct a simple equilibrium model of credit in which the presence of financial intermediaries is explicitly accounted for. Drawing on a literature that emerged in the late 1970s, the function of financial intermediaries is justified by the existence of an asymmetric information problem between borrowers and lenders (See Freixas and Rochet, 2008, chp1). In this model, the moral hazard generated by the possibility for borrowers to choose to undertake an inefficient project means that they may not be able to borrow externally without being monitored and investing some of their net worth into their project. The function of intermediaries is then to perform delegated monitoring on the behalf of outside investors[1]. In addition, intermediaries' creditors also face a moral hazard problem in that intermediaries may lack the proper incentive to perform the costly monitoring activity for which they are paid at a premium over the opportunity cost of funds in the economy. As a result, they also need to have some stake in the problem. This structure of the problem is flexible enough to accomodate two possible interpretations of financial intermediaries in this model; in the first one (intermediation case), they can be seen as banks financing loans granted to firms (borrowers) by using deposits from outside investors (households) and their own equity, whereas in the second one (certification case), they can be seen as venture capitalists or investment banks certifying the soundness of a project by investing some of their own funds in it, therefore enabling firms to attract outside investors funds as well.

This paper contributed to the literature by defining two types of capital constraints, one affecting borrowers (demand factor) and one affecting intermediary capital (supply factor) and distinguishing between bank and market debt, thus introducing a role for the distribution of capital across three types of agent (firm-borrowers, intermediaries and investors). As in

[1] This function of of intermediaries was already emphasized by Diamond (1984).

Bernanke and Gertler (1989), a shock that redistributes capital[2] between agents can affect investment and interest rates.

The model can account for some of the facts observed during a recession, when one type of capital becomes relatively scarce. In particular, it is able to explain how the health of the banking sector may have an impact on bank lending in addition to borrowers' balance sheet conditions. For example, a key feature of the credit crunch that hit the US economy in 1990-1991 is the persistent fall in bank lending. It is still an open debate how much of this fall is due to weaknesses in aggregate demand and how much is due to weaknesses in the financial sector (Jimenez et. al., 2012). An interesting fact pointing in the direction of a bank lending channel is the correlation between bank lending and banks' capital-asset ratios[3]. These ratios are a key concern of regulatory institutions and this model goes some way to provide an explanation for what would be their behaviour in the absence of regulatory constraints. The model also provides some intuition about the determinants of interest spreads.

Here are some of the key observed relationships the model, or some variant of the model, is able to explain.

1. A firm's net worth determines both its debt capacity and its source of financing. In particular, firms with substantial net worth prefer cheaper, less information-intensive (asset-backed) finance.

2. Highly leveraged firms (of which small firms may be a good proxy) favour more bank finance (monitoring).

3. Capital-poor firms are more affected by a fall in monitoring capital.

4. Credit crunches are associated with an increase in the spread between the cost of intermediated finance (for example bank loan rates) and the cost of market finance (for example deposit rates or bond rates).

5. In general, changes in this spread depend on changes in the relative amounts of the different types of capital.

6. An increase in monitoring capital relative to firm capital leads to more intensive monitoring in lending.

7. The efficient, market-determined capital adequacy ratios are procyclical.

3.1.1 The basic model

There are three types of agents holding different amounts of capital: firms, financial intemediaries and investors. Firms and financial intermediaries are distinguished by havinga unique access to an investment technology and a monitoring technology respectively. All agents are

[2] In the reminder of this section, we will refer to these three types of capital interchangeably as firm capital or net worth, monitoring or informed or financial intermediaries' capital and uninformed or investors' capital respectively.

[3] See Sharpe (1995).

risk-neutral and protected by limited liability. There are two markets, one for informed capital (bank loan for example) and one for uninformed capital (direct finance). The rates of interest that clear the two markets are respectively β and γ. The contracting relationship between agents unfolds over two periods. In the first period, they draw up the financial contracts which determine firms' investment decision. In the second period, once each firm's investment project is realized, agents settle their claims to its gross return.

The paper considers two models in which the nature of the investment project differs. In the fixed investment scale (FIS) model, the size of the investment is fixed and identical across firm, whereas in the variable investment scale (VIS) project, it is allowed to vary. These two models can be used to explain different facts (see section 6) and the behaviour of different variables. We describe and systematically compare them both.

The real sector

Firms There is a continuum of firms using the same technology but differing in their initial amount of capital (or net worth or assets) A. The capital asset can be cash or the real collateral value of any other type of asset. However its value cannot be influenced by the market interest rate. The cumulative distribution function of assets across firms is denoted $G(A)$. The aggregate amount of firm capital is $K_f = \int AdG(A)$.

Investment projects

FIS In period 1, each firm is endowed with one investment project which costs $I > 0$. If $A < I$, the firm will seek $I - A$ in external funds. In period 2, all agents are able to observe the financial return to the investment project which is either 0 (failure) or R (success).

VIS Investment can be undertaken at any scale $I(A)$. In case of success, the gross return becomes RI so that R is now the gross return per unit of investment and the production function exhibits constant returns to scale.

Moral hazard problem 1 Entrepreneurs can privately choose between two versions of the project, differing in their probability of success and the level of private benefits they can gain from it. The 'good' project has probability of success p_H and the 'bad' project has probability of success p_L, with $\Delta p = p_H - p_L > 0$. Private benefits constitute an additional source of revenue for entrepreneurs only when they choose the bad project and are not observed by the other agents. Monitoring by intermediaries (see below) reduces the level of private benefit generated by the 'bad' project without affecting its probability of success p_L.

FIS Without monitoring, the private benefit is B and with monitoring it is b, such that $0 < b < B$. Finally, we assume that only the 'good' project is efficient:

$$p_H R - \gamma I > 0 > p_L R - \gamma I + B \tag{3.1}$$

γ is the interest rate charged by investors and represents the opportunity cost of funds in this economy. The second inequality says that the total expected return (including entrepreneurs' private benefit) to the bad project is inferior to its opportunity cost. The first inequality says that the reverse holds for the good project and implies that firms' internal rate of return is higher than the cost of external finance[4].

Given that firms will not choose the bad project if their expected return[5] is less than the opportunity cost of funds ($p_L R_f + B \geq \gamma A$), (1) implies that:

$$p_L R + B < \gamma I$$

$$p_L R_u + p_L R_f + B < \gamma(I - A) + \gamma A$$

$$p_L R_u - \gamma(I - A) < \gamma A - (p_L R_f + B) \leq 0$$

$$p_L R_u - \gamma(I - A) < 0$$

If firms choose the bad project, investors' expected return to lending is less than their opportunity cost. Therefore they will not finance firms unless they are sure the good project will be undertaken (necessary condition for obtaining external funds) and any contract between firms and investors should ensure that firms do not have the incentive to choose the bad project. This implies that the existence of a debt contract guarantees an efficient outcome in the sense that a reallocation of funds between investors and firms would not lead to a Pareto improvement.

VIS Firms' private benefits are proportional to the investment size, so that B becomes BI and b becomes bI. Assumption (1) becomes:

$$p_H R - \gamma > 0 > p_L R - \gamma + B \tag{3.2}$$

Financial intermediaries

The function of intermediaries Financial intermediaries (FI) monitor firms and alleviate the moral hazard problem. Monitoring reduces entrepreneurs' private benefit in the event they choose the 'bad' project from B to b. There are many intermediaries, each with the physical capacity to monitor an arbitrary number of firms fully. It is assumed that the returns to the projects financed by each FI are perfectly correlated (It is not possible to diversify idiosyncratic risks so that FIs need to have a stake in the projects). In addition, monitoring

[4] This is a standard assumption in the literature on financial frictions due to problems of asymmetric information.

[5] Here we make use of the following notation: R_f is firms' gross return to the project and R_u is investors' gross return to lending such that $R_f + R_u = R$.

(or FI's) capital, K_m, is assumed to be relatively scarce and monitoring to be socially valuable and/or individually desirable[6].

Moral hazard problem 2

FIS Monitoring costs $c > 0$ to FIs. Whether FIs actually perform monitoring is not observed by investors. The equilibrium rate of return on intermediary capital (or bank lending) is β. Assuming the FI sector is not too concentrated, FIs take β as given. However, to account for the positive cost of monitoring, β has to be superior to γ (we show that this is the case later). Hence the moral hazard problem.

VIS The monitoring cost becomes cI.

Investors

Individual investors are uninformed (they do not have access to the monitoring technology). Because they are small, they take the equilibrium rate of return on uninformed capital, γ, as given. We assume some degree of elasticity in the supply of uninformed capital, $S(\gamma)$ (with reciprocal $\gamma(K_u)$) so that $S(\gamma)$ is upward-sloping in the (γ, S) space and the total amount of uninformed capital, K_u, is endogenous. Firms with excess capital invest it in the open market as uninformed capital.

The credit market equilibrium model

The basic model therefore features seven parameters (six in the VIS case), three endogenous variables and three possible sources of exogenous shock (see Table 3.1). Negative shocks to FIs' capital (K_m), firms' capital (K_f) and to investors' savings (K_u) are respectively called a *credit crunch*, a *collateral squeeze* and a *savings squeeze*. Aggregate investment is $K = K_f + K_m + K_u$ and is endogenous because uninformed capital K_u is endogenous.

parameters	p_H, p_L, B, b, c, R, I(FIS)
endogenous variables	β, γ (and $S(\gamma)$), K
exogenous variables/shocks	K_f, K_m, K_u

Table 3.1: A credit market equilibrium model

The analysis follows two steps. First, we derive firms' optimal investment decision as a function of β and γ. A financial contract between the three types of agents specifies (1) the amount each agents will contribute to the firm's investment project ($I_f + I_m + I_u = I$) and (2) their respective share of the gross return to the project ($R_f + R_m + R_u = R$) (see Table 3.2). In the FIS case, firms will invest if the value of their capital assets is above the threshold values $\overline{A}(\gamma)$ for direct finance and $\underline{A}(\gamma, \beta)$ for indirect finance. In the VIS case, the solution to the

[6]The exact conditions for these assumptions to hold are derived later.

contracting problem yields the optimal inverse leverage ratio which applies to all firms, $A_1(\gamma, \beta)$.

	Firm	FI	Investor
Amount of own capital invested in project	I_f	I_m	I_u
Gross return	R_f	R_m	R_u

Table 3.2: The terms of the contract

Second, credit market clearing conditions determine the equilibrium values of γ and β as well as the aggregate level of investment, K.

3.1.2 Fixed investment scale

Direct finance

This section specifies the terms under which a firm will borrow from uninformed investors only.

One optimal contract To derive the form of the optimal contract, we start by making the following working hypothesis. The firm invests all its funds A (this will not be necessary in equilibrium) in the project and the investor $I_u = I - A$. If the investment fails, neither party is paid anything. If the investment succeeds, the firm is paid $R_f > 0$ and the investor is paid $R_u > 0$ so that $R_f + R_u = R$.

Firms' incentive constraint For firms to choose the good project, it must be that:

$$p_H R_f \geq p_L R_f + B$$

This implies the following lower bound on the firm's share of the gross return:

$$R_f \geq \frac{B}{\Delta p} \tag{3.3}$$

This condition implies that the maximum expected income that can be promised investors without destroying firms' incentive to behave diligently (pledgeable expected income) is equal to:

$$p_H \left[R - \frac{B}{\Delta p} \right]$$

A necessary and sufficient condition for direct finance Investors' expected return from lending cannot be less than their opportunity cost of funds:

$$\gamma I_u \leq p_H R_u \tag{3.4}$$

Firms' investment decision To summarize firms' problem, they will invest if the following conditions hold:

$$A + I_u \geq I$$
$$R_f \geq \frac{B}{\Delta p}$$
$$p_H R_u \geq \gamma I_u$$
$$R \geq R_f + R_u$$

This program defines the minimum value of initial capital that firms must hold in order to obtain direct finance:

$$A \geq I - \frac{p_H}{\gamma}\left[R - \frac{B}{\Delta p}\right] = \overline{A}(\gamma) \tag{3.5}$$

$\overline{A}(\gamma)$ is increasing in γ: $\partial\overline{A}(\gamma)/\partial\gamma > 0$. As their opportunity cost of funds increases, investors ask for a higher gross return R_u for a given stake in the investment project I_u. Because at the optimum, firms' incentive constraint binds, R_u cannot be increased without destroying firms' incentive to behave diligently. As a result, I_u must decrease and firms must increase their stake in the project A. Finally, we note that firms with $A > \overline{A}(\gamma)$ are indifferent between investing the difference in the firm or in the market for uninformed capital.

Additional assumption 1 To rule out the case where $\overline{A}(\gamma) < 0$, we assume that the total surplus from a project is less than the minimum share a firm must be paid to behave diligently:

$$p_H R - \gamma I < \frac{p_H B}{\Delta p} \tag{3.6}$$

The total surplus generated by the project is equal to the expected gross return minus its opportunity cost. For the investment to be undertaken through direct finance, firms must choose the good project and to ensure they do so, they must be paid a minimum amount (defined in (3.5)) that is larger than the total surplus. This condition implies that the expected gross return to the investment project can be redistributed efficiently only if firms contribute a non-negative amount towards the cost of investment[7]. Efficiency here is not defined by total surplus maximization, but by the resolution of the agency problem.

Indirect finance

If $A < \overline{A}(\gamma)$, firms may still be able to obtain external funding if they accept to be monitored by FIs. This section specifies the terms under which a firm will be able to borrow from both informed and uninformed investors.

One optimal contract Assume the firm invests all its assets A, FIs contribute I_m and investors I_u. If the investment fails, no party is paid anything. If the investment succeeds, the

[7]This condition can be reformulated as follows: rewriting (3.6) as $\gamma I > p_H[R - B/\Delta p]$, it means that the total market value of the investment is superior to the pledgeable expected income so that investors will not contribute to the full amount of the investment cost.

firm is paid $R_f > 0$, the FI, $R_m > 0$, and the investor is paid $R_u > 0$ so that $R_f + R_m + R_u = R$.

Firms' incentive constraint For firms to choose the good project, it must be that:

$$p_H R_f \geq p_L R_f + b$$

This implies the following lower bound on the firm's share of the gross return:

$$R_f \geq \frac{b}{\Delta p} \tag{3.7}$$

FIs' incentive constraint For FIs to monitor firms, it must be that their expected gross return in the good case is higher than their expected return in the bad case:

$$p_H R_m - c \geq p_L R_m$$

It says that FIs' expected return when the good project is undertaken minus the monitoring cost incurred must be higher than their expected return when they do not monitor and firms choose the bad project[8]. This implies the following lower bound on the FI's share of the gross return:

$$R_m \geq \frac{c}{\Delta p} \tag{3.8}$$

The FIs' constraint also implies $p_H R_m - c > 0$ so that FIs earn a positive net return to monitoring. Assuming monitoring capital is scarce, FIs make a strictly positive profit[9].

A necessary and sufficient condition for indirect finance The expected gross return to FIs' lending cannot be less than the their opportunity cost of funds:

$$p_H R_m \geq \beta I_m \tag{3.9}$$

Because of the monitoring cost incurred by FIs, β must be high enough to make the intermediary prefer monitoring to investing in the market for uninformed capital. This imposes a lower bound on β:

$$\beta I_m - c \geq \gamma I_m \quad \rightarrow \quad \beta \geq \frac{c}{I_m} + \gamma > \gamma$$

As a result, firms will always prefer uninformed to informed capital and will demand precisely the minimum level of informed capital required. In equilibrium, both inequalities (3.8) and (3.9) are binding. Combining them both we obtain the minimum acceptable rate of return on informed capital, $\underline{\beta}$, which is defined when the last constraint holds with equality:

$$\underline{\beta} = \frac{\gamma p_H}{p_L}$$

[8]Because $b < B$, it will always be optimal for firms to choose the bad project if firms choose not to monitor them

[9]According to Holmstrom and Tirole [40], an increase in monitoring capital, that could be interpreted as an increase in competition in the financial intermediary sector, would reduce this surplus.

Firms' investment decision Firms will invest if the following conditions hold:

$$A + I_m + I_u \geq I$$

$$R_f \geq \frac{b}{\Delta p}$$

$$p_H R_u \geq \gamma I_u$$

$$R_m \geq \frac{c}{\Delta p}$$

$$p_m R_m \geq \beta I_m$$

$$R \geq R_f + R_m + R_u$$

This program defines the minimum amount of firm's capital that firms must hold in order to obtain indirect finance:

$$A \geq I - I_m(\beta) - \frac{p_H}{\gamma}\left[R - \frac{b+c}{\Delta p}\right] = \underline{A}(\gamma, \beta) \tag{3.10}$$

(3.8) and (3.9) define the amount of capital FIs invest in the project: $I_m(\beta) = p_H c/\beta\Delta p$ is decreasing in β. $\underline{A}(\gamma, \beta)$ is increasing in both γ and β: $\partial\underline{A}(\gamma, \beta)/\partial\gamma > 0$ and $\partial\underline{A}(\gamma, \beta)/\partial\beta > 0$.

Additional assumption 2 A necessary and sufficient condition for financial intermediation to take place and be socially useful is given by:

$$\underline{A}(\gamma, \beta) > \overline{A}(\gamma) \quad \rightarrow \quad p_H(B - b) > c\Delta p$$

Additional assumption 3 To rule out the possibility of an equilibrium with monitoring where FIs do not need to invest their own capital into the project we need to assume that $b + c > B$.

Thus, depending on their net worth, firms will choose either one of these three options: (1) if $A \geq \overline{A}(\gamma)$ they choose market finance only and may invest their surplus either in their investment project or in the market for uninformed capital, (2) if $\underline{A}(\gamma, \beta) \leq A < \overline{A}(\gamma)$ they will seek both intermediated finance and market finance and (3) if $A < \underline{A}(\gamma, \beta)$ they obtain no external funding and invest their net worth in the market for uninformed capital.

Equilibrium in the credit market

Equilibrium in the monitoring market Taking γ as given, the equilibrium value of β is such that the total stock of monitoring capital is equal to the total demand for monitoring capital:

$$K_m = D_m(\gamma, \beta) = \left[G(\overline{A}(\gamma)) - G(\underline{A}(\gamma, \beta))\right] I_m(\beta) \tag{3.11}$$

To avoid the case of a corner solution, it is assumed that monitoring capital is sufficiently rare so that the equilibrium value of β is strictly superior to the minimum acceptable rate of return $\underline{\beta}$. Since $\partial D_m/\partial\beta < 0$, for each γ, there is a unique value of β that clears the market. However the effect of a change in γ on the demand for monitoring capital is ambiguous. On one hand, an increase in γ raises $G(\overline{A}(\gamma))$ and therefore the number of firms that can access market finance directly, thus increasing the number of firms that seeking intermediated finance. On the other hand, it raises $G(\underline{A}(\gamma,\beta))$ or the number of firms that are squeezed out of the market for external finance altogether, thus reducing the number of firms asking for monitoring capital. As a result, the sign of $\partial D_m/\partial\gamma$ depends on the exact form of the distribution $G(A)$.

Equilibrium in the market for uninformed capital Taking β as given, the equilibrium value of γ is such that the total supply of uninformed capital is equal to the total demand for ininformed capital:

$$S(\gamma) = D_u(\gamma,\beta) = \int_{\underline{A}(\gamma,\beta)}^{\overline{A}(\gamma)} [I - A - I_m(\beta)]\, dG(A) + \int_{\overline{A}(\gamma)}^{\infty} [I - A]\, dG(A) \tag{3.12}$$

Since $\partial D_m/\partial\gamma < 0$, for each β, there is a unique value of γ that clears the market. However the effect of a change in β on the demand for uninformed capital is ambiguous. On one hand, an increase in β increases $\underline{A}(\gamma,\beta)$ and the number of firms able to invest altogether decreases, thus reducing the demand for uninformed capital. On the other hand, it decrease $I_m(\beta)$, the amount invested in each project by FIs, thus requiring a larger contribution of uninformed capital in the projects that are financed. As a result, the sign of $\partial D_m/\partial\beta$ depends on the exact form of the distribution $G(A)$.

Equilibrium in the market for external finance Conditions (3.11) and (3.12) can be combined into one equilibrium condition where the total demand for external capital by firms is equated to the total supply of monitoring and uninformed capital:

$$\int_{\underline{A}(\gamma,\beta)}^{\infty} [I - A]\, dG(A) = S(\gamma) + K_m \tag{3.13}$$

Changes in the supply of capital

Result 1

> 'In either type of capital squeeze, aggregate investment will go down, and $\underline{A}(\gamma,\beta)$ will increase.' (p 678)

To prove this proposition, assume first that $\underline{A}(\gamma,\beta)$ decreases so that aggregate investment goes up. This means that the supply of uninformed capital increases. This is only possible if the rate of return on uninformed capital γ increases. When γ increases, $\overline{A}(\gamma)$ increases as well. Since $\underline{A}(\gamma,\beta)$ falls and $\overline{A}(\gamma)$ rises, a larger number of firms are eligible for monitoring capital. Because the supply of monitoring capital is fixed (and/or has just fallen), FIs' contribution to each firms' investment project $I_m(\beta)$ must fall, or equivalently, β must rise. But with both γ and β rising, $\underline{A}(\gamma, \beta)$ should increase. This contradicts our initial assumption and proves Result 1 [10].

In a recession, we might expect R and p_H to decline. This would tend to reduce $\underline{A}(\gamma, \beta)$ as well.

3.1.3 Variable investment scale

The firm's program

Let $U(A)$ be the firms' payoff which is equal to its gross return to the investment project and on other assets. Then the firm's problem is to choose the level of investment (or equivalently its leverage ratio defined as I/A) subject to the same set of constraints as before:

$$max\ U(A) = p_H RI - p_H R_m - p_H R_u + \gamma(A - I_f)$$

Subject to

$$A \geq I_f$$

$$A + I_m + I_u \geq I$$

$$R_f \geq \frac{bI}{\Delta p}$$

$$p_H R_u \geq \gamma I_u$$

$$R_m \geq \frac{cI}{\Delta p}$$

$$p_m R_m \geq \beta I_m$$

$$RI \geq R_f + R_m + R_u$$

This program defines firms' highest sustainable leverage ratio as:

$$\frac{I}{A} \leq \frac{1}{A_1(\gamma, \beta)} \tag{3.14}$$

$$A_1(\gamma, \beta) = 1 - I_m(\beta) - \frac{p_H}{\gamma}\left[R - \frac{b+c}{\Delta p}\right] < 1$$

$A_1(\gamma, \beta)$ is the minimum amount of firm capital needed to undertake an investment of size 1. The lower it is, the higher the firm's leverage.

Additional assumption 4 In equilibrium, γ and β must be such that $A_1(\gamma, \beta) > 0$ otherwise firms' optimal investment level would be unbounded.

[10] This demonstration also proves that the equilibrium defined by $\underline{A}(\gamma, \beta)$ and $\overline{A}(\gamma)$ must be unique.

Firms' maximum payoff is given by:

$$U(A) = \frac{p_H b}{\Delta p} I(A) \quad where \quad I(A) = A/A_1(\gamma, \beta)$$

It is increasing in the level of investment and in the leverage ratio. It is optimal for firms to invest up to the maximum possible leverage ratio and all the constraints in firms' programs bind in equilibrium. The net value of firms' leverage is given by the difference between firms' payoffs when borrowing externally and investing in their project (the internal rate of return) and the opportunity cost of funds (the external rate of return):

$$\left[\frac{p_H b}{\Delta p A_1(\gamma, \beta)} - \gamma \right] A \tag{3.15}$$

The term in bracket (and therefore the whole expression) is positive if the following condition (evaluated at $\underline{\beta}$) holds:

$$p_H R - \gamma > c \quad \leftrightarrow \quad \frac{c}{p_H R - \gamma} < 1$$

Additionnal assumption 5 For monitoring to be of value, the difference between firms' maximum payoffs when using monitoring capital and when using direct finance only should be positive (evaluated at $\underline{\beta}$)[11]:

$$1 > \frac{B-b}{B} > \frac{c}{p_H R - \gamma}$$

This condition ensures that equation (3.15) is positive.

Equilibrium in the capital markets

Equilibrium in the market for uninformed capital Let $\gamma = \gamma(K_u)$ be the inverse supply function of uninformed capital $(S^{-1}(\gamma))$ which characterizes γ as an increasing function of K_u. The equilibrium rate of return γ is such that total expected pledgeable income is equal to investors' total opportunity cost for uninformed funds:

$$p_H K \left[R - \frac{b+c}{\Delta p} \right] = \gamma(K_u) K_u \tag{3.16}$$

$$\gamma = p_H \frac{K}{K_u} \left[R - \frac{b+c}{\Delta p} \right] \tag{3.17}$$

The equilibrium value of γ is larger than the expected pledgeable income per unit of investment $p_H[R - (b+c)/\Delta p]$.

Equilibrium in the market for informed capital The equilibrium rate of return β is such that the total expected return to informed capital is equal to the total opportunity cost

[11] This expression differs from the one given by Holmstroma and Tirole: $c(p_H\gamma - p_L)/\Delta p < (B-b)/B$. To verify.

for informed funds:

$$p_H \frac{cK}{\Delta p} = \beta K_m \tag{3.18}$$

$$\beta = \frac{p_H c}{\Delta p} \frac{K}{K_m} \tag{3.19}$$

Equations (3.17) and (3.19) imply that the equilibrium rates of return depend crucially on the relative shares of uninformed and informed capital in aggregate investment respectively.

Changes in the supply of capital

Solvency ratios Firms' solvency ratio is defined as the inverse of firms' leverage ratio, which is the same as the aggregate leverage ratio in the economy: $r_f = K_f/K$. The definition of the FIs' solvency ratio, $r_m = K_m/(K_m + K_u)$ only makes sense in the context of the 'intermediation' case, whereby FIs represent commercial banks financing their loans to firms in part by taking deposits from investors. The solvency ratio is the share of equity in the banks' total liabilities. According to this view the equilibrium outcome also determines how much deposits (K_u) banks will be able to attract.

Result 2 The results presented in Table 3.3 follow directly from equations (3.16)-(3.19). Following any type of capital squeeze, both K and K_u fall.

	credit crunch ($K_m \downarrow$)	collateral squeeze ($K_f \downarrow$)	savings squeeze ($S(\gamma) \downarrow$)
β	$\uparrow$	$\downarrow$	$\downarrow$
γ	$\downarrow$	$\downarrow$	$\uparrow$
r_m	$\downarrow$	$\uparrow$	$\uparrow$
r_f	$\uparrow$	$\downarrow$	$\uparrow$

Table 3.3: The effect of different types of capital squeeze on key financial variables

These outcomes depend crucially on the way exogenous shocks affect the relative scarcity of the three types of capital. In particular, since firms' net worth and monitoring capital are assumed to be exogenously given, changes in the supply of uninformed capital are crucial. As in Holmstrom and Tirole (1997), we take the example of a credit crunch. From (3.16), the fall in K_m leads to a fall in the right-hand side of the equation, which can only happen if both γ and K_u decrease. As a result, K_f/K_u rises and K_m/K_u falls. K_u decreases proportionally less than K_m because the fall in γ necessary to accomodate the fall in K_u renders uninformed capital cheaper. K/K_m increases, so that the share of monitoring capital in total investment falls. The demand for monitoring does not fall by as much as the supply of monitoring capital, which is only consistent with a rise in β (3.19). As the supply of external funds goes down, firms put up more of their own asset in each unit of the investment project, and their solvency ratio rises. FIs are able to attract more deposit for each unit of equity that they own and their solvency ratio goes down.

3.1.4 Two extensions of the basic model

Endogenous monitoring

FIS This section considers allowing both the intensity of monitoring c and the private benefit associated with the bad project b to vary. Because capital-rich firms will want to reduce the amount of monitoring capital needed (to substitute it for cheaper uninformed capital) by reducing monitoring costs, the intensity of monitoring is continuously decreasing with the level of firms' assets. The main point here is that higher monitoring costs require more monitoring capital for FIs to retain the right incentives.

VIS The level of monitoring intensity depends on the relative amounts of firms' and informed capital. The main point here is that a higher amount of monitoring capital relative to firm capital leads to a higher intensity of monitoring and allows for a higher leverage of firms and therefore a higher level of aggregate investment. This could be achieved by transfering capital from firms to FIs, while keeping the aggregate amount of capital in the economy constant. For this reason, from the policy point of view, 'it may be more efficient to subsidize intermediaries than to subsidize firms' (p687).

Decreasing returns to scale

Finally, the VIS model can be modified so that the investment project is subject to decreasing returns to scale ($RI \to R(I)$). In this case, the optimal level of investment $I(A)$ becomes an increasing and concave function of A and it will not be optimal for firms to increase leverage indefinitely. In other words, this assumption introduces a discontinuity in the form of a threshold level of investment above which it will not be optimal for firms to invest more. As a result, optimal leverage will be lower for capital-rich firms that reach this level. The effect of a credit crunch for those firms is ambiguous and depends on the shape of $R(I)$.

3.1.5 Fixed investment scale versus variable investment scale

In Table (3.4), we summarize the key results obtained in the two variants of the model to emphasize their complementarity. In particular, it can be seen that of the seven relationships the authors propose to explain and listed in the introduction, the FIS model is more appropriate to explain propositions 1 and 3, as well as proposition 2 when endogenous monitoring is added. It can also account for the 'flight-to-quality' phenomenon. The VIS model can explain only the first part of proposition 1, propositions 4, 5 and 7, as well as proposition 6 when endogenous monitoring is added. In general, the FIS model is more appropriate to explain facts that can only be explained at the micro level, such as the flight to quality whereas the VIS case is more appropriate to explain macroeconomic phenomena such as the determination of interest rates.

	Fixed investment scale	**Variable investment scale**
Optimal contract	$\overline{A}(\gamma) = I - p_H/\gamma[R - B/\Delta p]$ $\underline{A}(\gamma,\beta) = I - I_m(\beta) - p_H/\gamma[R - (b+c)/\Delta p]$	$A_1(\gamma,\beta) = 1 - I_m(\beta) - p_H/\gamma[R - (b+c)/\Delta p]$
Aggregate investment	$K = I \int_{\underline{A}(\gamma,\beta)}^{\infty} dG(A)$	$K = K_f/A_1(\gamma,\beta)$
Credit rationing	Firms with $A < \underline{A}(\gamma,\beta)$ cannot raise external funds	All firms can raise some funds
Leverage	Varies across firms (increases with A)	Identical for all firms: $1/A_1(\gamma,\beta)$
Which firms ask for bank capital?	Only firms with $\underline{A}(\gamma,\beta) < A < \overline{A}(\gamma)$	All firms (allows for higher leverage)
Scale economies in monitoring	Yes	No (monitoring cost increases with size)
Effect of exogenous shocks	$K \downarrow (K_u \downarrow)$ $\underline{A}(\gamma,\beta) \uparrow$ (capital-poor firms hit first) At least one interest rate $\uparrow$ The effect on the other is ambiguous	$K \downarrow (K_u \downarrow)$ Credit crunch: $\gamma \downarrow$ and $\beta \uparrow$ Collateral squeeze: $\gamma \downarrow$ and $\beta \downarrow$ Saving squeeze: $\gamma \uparrow$ and $\beta \downarrow$

Table 3.4: Fixed investment scale vs variable investment scale

3.1.6 Conclusion

This model shows how the existence of moral hazard problems between agents in the financial system may provide some intuition about the organization of credit markets and the way they may affect aggregate investment. Contrary to the models discussed in chapter 1, intermediaries cannot (or do not want to) diversify firms' idiosyncratic risks. Having some degree of correlation between individual firms' returns implies that financial intermediaries need to put up some of their own capital into firms' investment projects. Subsequently, the relative scarcity of FIs' capital play a key role in the way credit markets react to exogenous shocks. A second important difference is the shift of focus from asset prices to the distribution of capital across the three types of agents. As in Kiyotaki and Moore (1997), interest rates do not reflect fundamentals, which may only provide the bounds within which they fluctuate according to the relative scarcity of capital.

This provides several new insights policy analysis. A change in the policy rate may have effects on investment by affects the relative amount of capital held by each type of agent in the economy. Regarding the debate over banks' solvency ratios, the fact that they appear to be procyclical in a credit crunch (and countercyclical otherwise), implies that a fall in FIs' solvency ratios during a recession may be an efficient way the economy has to cope with adverse events.

3.2 Gertler and Kiyotaki (2011)

The model developed in Gertler and Karadi (2011) has a similar structure as in BGG, except that the financial friction, instead of constraining borrowers, here applies to financial intermediaries from which they borrow the funds necessary to invest. The agency problem takes the form of a moral hazard problem giving rise to a spread between the loan rate and the deposit rate. FIs' leverage is endogenous and their net worth becomes an important state variable. Gertler and Kiyotaki (2011) is a generalisation of the model in Gertler and Karadi (2011) to allow for liquidity shocks and an interbank market exhibiting the same type of financial frictions as in the retail (deposit) market.

In Gertler and Karadi (2011), the financial friction is embedded into a standard NK model, allowing the study of both conventional and unconventional monetary policies as well as a reconsideration of optimal monetary policy. In Gertler and Kiyotaki (2011), the financial friction is embedded into an RBC model and the focus of the paper is on the real effects of credit policies of the unconventional type carried on by central banks during the recent financial crisis. This second model captures the idea of interconnectedness within the financial system, which is brought by the distinction between the interbank and the retail market, and examines its implication for the propagation of shocks to the real economy. Thus the emphasis in this paper is on monetary policy in time of crisis[12], triggered by an exogenous fall in the quality of capital assets[13].

[12]It is possible to imagine that the incentive constraint faced by financial intermediaries in this model starts binding in time of crisis, when there is a sharp fall in financial intermediariEs' net worth.

[13]In GK1, the authors also investigate the effect of the news shock of an expected fall in the quality of assets

3.2.1 The basic setup

The representative household For credit market frictions to matter, there must be some borrowing and lending in equilibrium, and thus some degree of heterogeneity between agents. Gertler and Kiyotaki (2011) adopt the consolidated family approach of the representative household. Each household is composed of a continuum of members of mass unity of which bankers constitute a constant fraction f of them, the remaining ones being workers. Bankers exit with probability $(1-\sigma)$ in each period and are automatically replaced by workers[14]. When they exit, bankers transfer their last period net worth to the household budget while new bankers receive a fraction $\xi/(1-\sigma)$ of the assets of exiting bankers.

The households' problem is to choose their level of consumption, labour and deposits to maximize their lifetime utility subject to their budget constraint. Because households own non-financial firms, capital producing firms and banks, the objective of these three entities will be to maximize the present discounted value of their respective streams of profits.

Islands The structure of the financial system is illustrated in Figure 3-3. The economy is composed of a continuum of islands on which non-financial firms and banks contract with each other. Firms' capital is fixed and they can only borrow from banks on the same island. However, both workers and bankers are perfectly mobile (they can move between islands form one period to the next). This assumption allows introducing a limited participation constraint in the market for firms' securities: firms can sell securities only to the banks on the same island. In addition, there are two perfectly competitive financial markets operating at the national level to which banks have access: one for retail deposits and one for wholesale (interbank) loans.

Non-financial firms Competitive non-financial firms produce the output good using a CRS production function. For each unit of capital asset used by firm, banks hold a state-contingent security (s_t^h) worth Q_t in their balance sheets. Firms' capital assets at the end of the period consist in current period capital net of depreciation and investment. Firms' capital depreciates at rate δ in each period. Investment opportunities are assumed to arrive randomly with probability π^i (i.i.d. across islands and time) on each island. Thus each island is indexed by $h = i, n$ in each period depending on whether they invest or not. Firms on investing islands buy capital goods from capital producing firms at price Q_t^i and finance it by issuing new securities. The return to firms' capital is denoted R_{kt}^h (it is island dependent because in equilibrium the price of securities Q_t^h is island dependent) and is determined is a similar way as in BGG as the sum of the marginal return to capital (dividend payment) and capital gain:

$$R_{kt+1}^{hh'} = \psi_{t+1} \frac{\alpha Y_{t+1}/K_{t+1} + (1-\delta)Q_{t+1}^{h'}}{Q_t^h}$$

ψ_t is an exogenous shock to capital asset quality. h' is the type (state) of the island in period t+1 and is unknown to all agents at the end of period t when securities are sold so that the

and show that it has similar effect, even though it may not materialize.

[14] In each period, $(1-\sigma)f$ bankers become workers and the same number of workers become bankers.

return is defined conditional on this information in period t. Since this expression also defines the state-contingent return paid on each security, non-financial firms make no profit.

Capital-producing firms New capital goods are produced using the final good as input and exhibit decreasing returns to scale in the short run: as the rate of growth of investment increases, production increases at a decreasing rate. Capital-good producers choose the level of aggregate investment I_t to maximize their expected stream of future profits. The first-order condition defines the supply-curve for new capital goods which is increasing in the (I, Q^i) space. The demand for capital goods is determined by banks' optimal value of asset holdings.

Banks The financial intermediaries's sector encompasses both commercial and investment banks. At the end of each period, banks choose the volume of securities (s_t^h) they want to hold on the asset side of their balance sheets as well as the amount of short-term debt they will hold on the liability side of their balance sheet, either in the form of household deposit (d_t) or interbank loans (b_t), taking as given their respective interest rates R_{kt+1}^h, R_{t+1} and R_{bt+1}.

Assets	Liabilities
$Q_t^h s_t^h$	n_t^h
	b_t^h
	d_t

Table 3.5: Banks' balance sheet

The flow-of-funds constraint they face directly follows from the structure of their balance sheet (see Table 3.5):

$$Q_t^h s_t^h = n_t^h + b_t^h + d_t \tag{3.20}$$

where n_t^h represents banks' net worth and depends on the previous period profits in the case of continuing bankers:

$$n_t^h = R_{kt}^h Q_{t-1} s_{t-1} - R_{bt} b_{t-1} - R_t d_{t-1} \tag{3.21}$$

and net transfers from households in the case of new bankers as described in Figure 3-1. The

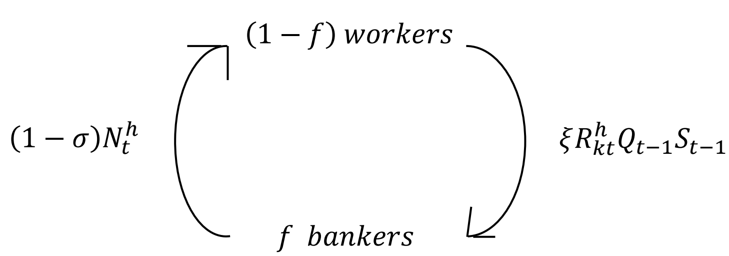


Figure 3-1: The consolidated family household: transfers of wealth between bankers and workers

expressions next to the arrows indicate the transfers between workers and bankers in each period. To facilitate aggregation, additional assumptions regarding the distribution of net worth across islands are made (described later) so that is becomes irrelevant; for these reasons, the arrows represent aggregate transfers.

Banks make their decisions sequentially, as described in Figure 3-2. Based on the expected issuance of new securities by non-financial firms on their island, they decide how much deposits to take on (retail market). Next, the random arrival of investment opportunities means that some banks will have surplus funds and some will have deficit funds (liquidity shock) that they will want to trade in the interbank market, at which point they determine their demand schedules for securities in each island.

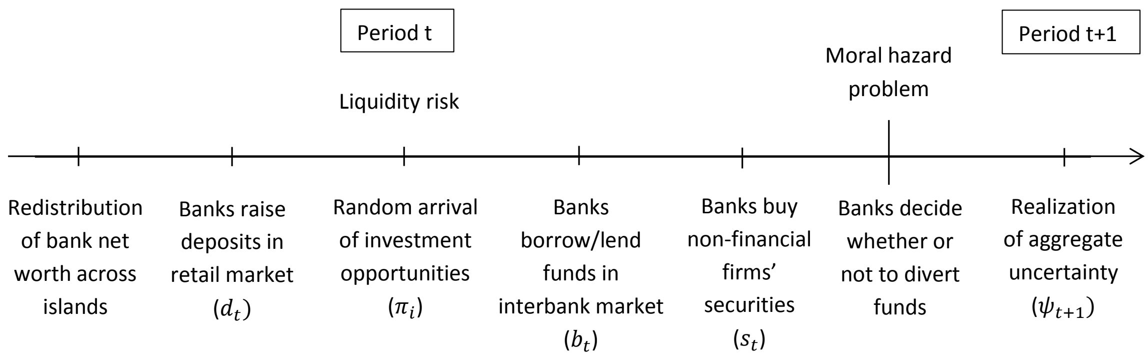


Figure 3-2: Sequence of events

The objective of the bank is to maximize the present discounted value of the future transfer of net worth[15] (to the household) when they exit:

$$V_t = E_t \sum_{i=1}^{\infty} (1-\sigma)\sigma^{i-1}\Lambda_{t,t+i} n_{t+i}^h \tag{3.22}$$

where $\Lambda_{t,t+i}$ is the household's stochastic discount factor (determined via the consumption' Euler equation), $(1-\sigma)$ is the probability that the banker exits in period i and σ^{i-1} is the probability that the banker survived until period i.

3.2.2 Credit and the demand for capital goods

The agency problem Bankers' creditors face a moral hazard problem, the introduction of which serves several purposes: to underlie the strong connection between the presence of asymmetric information problems between borrowers and lenders and the function of financial intermediaries, to motivate the presence of a credit spread in equilibrium, to give a key role to financial intermediaries in the propagation of shocks to the real economy and in particular to some aggregates in banks' balance sheet and to endogenize this mechanism. This is obtained

[15]It also represents the market value of the bank' s portfolio of assets?

by assuming that banks are constrained in the amount of debt they can take on by their performance in the previous period. This idea is captured by the introduction of a state variable taking the form of banks' net worth (inside equity).

Moral hazard is introduced by assuming that bankers are able to divert a fraction of the assets they hold before production starts and, in doing so, default on their debt. First, we define divertable assets as:

$$Q_t^h s_t^h - \omega b_t^h$$

where ω represents the fraction of interbank borrowing that cannot be diverted[16]. Second, we assume that bankers may transfer a fraction θ of the divertable assets to the household, therefore defaulting on their debt and leaving their creditors with the remaining $(1-\theta)$. Third, creditors will require that the stake of bankers in their portfolio investment (n_t^h) be high enough so that the gain form diverting assets and not paying interest on debt may be offset by the loss from being able to recover only a fraction of their initial stake in the project (θn_t^h).

For bankers to decide not to divert funds, the maximum present discounted value of their portfolio of assets at the end of the period must be at least equal to the value of the the assets they would be able to divert. Let $V_t(s_t^h, b_t^h, d_t)$ be the value function such that:

$$V_t(s_t^h, b_t^h, d_t) = \max V_t$$

and satisfying the Bellman equation:

$$V_t(s_t^h, b_t^h, d_t) = \max E_t \Lambda_{t,t+1} \sum_{h'=i,n} \pi^{h'} \left\{ (1-\sigma) n_{t+1}^{h'} + \sigma V_{t+1}(s_{t+1}^{h'}, b_{t+1}^{h'}, d_{t+1}) \right\} \tag{3.23}$$

where h' denotes the state of the bank in the next period. Then, bankers' incentive constraint is defined as follows:

$$V_t(s_t^h, b_t^h, d_t) \geq \theta \left(Q_t^h s_t^h - \omega b_t^h \right) \tag{3.24}$$

It says that the maximum achievable present discount value of bankers' final net worth must be at least as large as the value of divertable assets at the end of time t.

With frictionless financial markets, banks maximizing their expected discounted terminal net worth in each period would do so by equating marginal gain from buying an additional unit of securities to its marginal cost and these choices would only depend on current period opportunities (and not on the previous period net worth). Assuming banks are constrained in the volume of debt they can take on, however, a spread may open between marginal gains and marginal costs. In addition, because (1) banks live for several periods, (2) they maximize their expected stream of profits and (3) their profits in one period may determine how much they will be able to invest in the next period, their (internal) marginal value of buying securities and their (internal) marginal costs of debt may not coincide with the current period market

[16] It is inversely related to the degree of friction in the interbank market. A frictionless interbank market would correspond to a value of ω equal to 1. Reciprocallly, a value of ω equal to 0 would mean that the financial friction applies symetrically in the two financial markets.

interest rates and will depend on the marginal value of next worth in the next period[17].

The general solution To solve banks' problem, we proceed in four steps. First we make a guess about the form of the value function and investigate how the presence of financial frictions affect the spreads between marginal value and marginal costs. Second, we solve for banks' optimal portfolio of securities. Third, using the method of undetermined coefficients and the Bellman equation (3.23), we derive the relationships between (internal) marginal gains and cost and market interest rates. Finally, aggregating over all banks, we can derive banks' aggregate demand for assets and their aggregate net worth. We consider the two extreme cases of perfect and imperfect interbank market.

Step 1 Suppose the value function is linear in s_t^h, b_t^h and d_t:

$$V_t(s_t^h, b_t^h, d_t) = \upsilon_{st} s_t^h - \upsilon_{bt} b_t^h - \upsilon_t d_t \tag{3.25}$$

where υ_{st}, υ_{bt} and υ_t represent respectively the marginal value of assets and the marginal costs of interbank loans and deposits. All three are time-varying and identical in both types of island (why? see footnote 11 p558).

Step 2 Using the Lagrangean approach, we solve for the optimal spreads between marginal value and marginal costs as a function of the parameters θ and ω and the lagrangean multipliers associated with bankers' incentive constraints in both investing and non-investing islands. The Lagrangean for this problem is:

$$L = \sum_{h=i,n} \pi^h \left\{ \upsilon_{st} s_t^h - \upsilon_{bt} b_t^h - \upsilon_t d_t + \lambda_t^h \left[\upsilon_{st} s_t^h - \upsilon_{bt} b_t^h - \upsilon_t d_t - \theta \left(Q_t^h s_t^h - \omega b_t^h \right) \right] \right\} \tag{3.26}$$

Substituting $(Q_t^h s_t^h - n_t^h - d_t)$ for b_t^h and solving for d_t and s_t^h, we obtain three optimality conditions:

$$\upsilon_{bt} - \upsilon_t + \overline{\lambda}_t \left(\upsilon_{bt} - \upsilon_t - \theta\omega \right) = 0$$

$$\left(\upsilon_{bt} - \upsilon_t \right) \left(1 + \overline{\lambda}_t \right) = \theta\omega\overline{\lambda}_t \tag{3.27}$$

where $\overline{\lambda}_t = \sum_{h=i,n} \pi^h \lambda_t^h$ (to account for the fact that d_t is chosen before the realization of the liquidity shock).

$$\upsilon_{st} - \upsilon_{bt} Q_t^h + \lambda_t^h \left[\upsilon_{st} - \upsilon_{bt} Q_t^h - \theta \left(Q_t^h - \omega Q_t^h \right) \right] = 0$$

$$\left(\frac{\upsilon_{st}}{Q_t^h} - \upsilon_{bt} \right) \left(1 + \lambda_t^h \right) = \theta \left(1 - \omega \right) \lambda_t^h \tag{3.28}$$

Finally, substituting $(Q_t^h s_t^h - n_t^h - b_t^h)$ for d_t, the incentive constraint becomes:

$$\upsilon_t n_t^h \geq \left[\theta - \left(\frac{\upsilon_{st}}{Q_t^h} - \upsilon_t \right) \right] Q_t^h s_t^h - \left[\theta\omega - \left(\upsilon_{bt} - \upsilon_t \right) \right] b_t^h \tag{3.29}$$

[17] These internal marginal values represent the shadow prices of financial assets to banks.

which holds with equality when $\lambda_t^h > 0$ and as a strict inequality when $\lambda_t^h = 0$.

Condition (3.27) describes the conditions under which the shadow cost of interbank borrowing exceeds the shadow cost of deposits and condition (3.28) describes the conditions under which the shadow value of assets in terms of final goods in each type of islands ($\frac{\upsilon_{st}}{Q_t^h}$) exceeds the shadow cost of interbank borrowing. We can summarize conditions (3.27) and (3.28) by this set of inequalities:

$$\frac{\upsilon_{st}}{Q_t^h} \geq \upsilon_{bt} \geq \upsilon_t$$

The first inequality holds strictly if the incentive constraint binds in island h and there is at least some degree of financial friction in the interbank market ($\omega < 1$). The second inequality holds strictly if the interbank market is relatively more efficient ($\omega > 0$) and the incentive constraint is expected to bind in at least one of the two types of islands ($\overline{\lambda}_t > 0$). Note that if the incentive constraint is not expected to bind in both types of islands, for example because banks' net worth is high enough, then there is no agency problem, and in equilibrium banks choose the volumes of deposits, securities and interbank loans that equalize the three marginal values (or shadow prices).

Condition (3.29) says that banks' shadow value of net worth must be at least as large as a weighted measure of asset values net of interbank borrowing which depends on the fraction of divertable assets that can be appropriated by bankers (θ), the relative degree of friction in the interbank market (ω) and the magnitude of the two spreads between shadow prices. This last optimality condition represents the endogenous balance sheet constraint. It holds tighter the higher θ (the fraction of divertable assets bankers may be able to abscond with) and the lower the excess marginal value of assets over deposits. In the case of investing islands ($b_t^h > 0$), it is tighter the higher the degree of financial frictions in the interbank market (the lower ω) and the higher the spread between the marginal costs of interbank loans and deposits. Since the reverse holds in non-investing islands we might expect the constraint to hold tighter in investing islands ($\lambda_t^i \geq \lambda_t^n$).

Finally, we need to consider the implications of these optimality conditions for the determination of asset prices in both islands. On investing islands, asset prices are determined jointly by the capital supply schedule of capital-producing firms and the asset demand schedule of banks whereas in non-investing island they are determined by the arbitrage condition between securities and interbank loans. From (3.28), we can see that the spread between the shadow value of a unit of assets in terms of the numeraire and the shadow cost of interbank loans in each types of island is increasing in the value of the lagrange multipliers and in the degree of friction in the interbank market. Thus in investing island where the lagrange multiplier is larger, we would expect this spread to be larger and by implication, the asset price to be lower. In non-investing islands, as long as the shadow value of assets exceeds the shadow cost of interbank loans, banks will demand securities rather than loans, thus boosting asset prices. Depending on their net worth (which itself depends in part on the current value of asset prices), they may be able to reduce the spread to zero, in which case the lagrange multiplier is reduced to zero as well and the constraint does not bind in those islands. Therefore, in the general case, we may expect asset prices in investing islands (where firms issue more securities relatively to banks net worth) to be lower than in non-investing islands. Furthermore, the higher hte degree of frictions in the interbank market, the higher asset price dispersion (volatility).

Step 4 We assume that at the beginning of each period, before each island' s type is known, some bankers on non-investing islands (in the previous period) are able to sell their securities in exchange for interbank loans held by the others and move with their assets to investing islands (in the previous period) in such a way that the ex-ante expected return per unit of net worth are equalized across islands[18]. As a result, the only reason why net worth differs across islands in each period is because the random arrival of investment opportunities affects current period's asset prices, which in turn affects asset returns R^h_{kt}. This assumption facilitates the following aggregation results (?). Aggregate net worth on each type of island is given by the sum of the net worth of continuing (N^h_{ot}) and new bankers (N^h_{yt}):

$$N^h_t = N^h_{ot} + N^h_{yt}$$

where the first term is defined as:

$$N^h_{ot} = \sigma\pi^h \left[R^h_{kt} Q_{t-1} S_{t-1} - R_{t-1} D_{t-1}\right] \tag{3.30}$$

where $D_{t-1} = \sum_{h=i,n} \left(Q^h_{t-1} S^h_{t-1} - N^h_{t-1}\right)$ and the second term is defined as:

$$N^h_{yt} = \xi\pi^h R^h_{kt} Q_{t-1} S_{t-1} \tag{3.31}$$

These give the following expression for banks' aggregate net worth:

$$N^h_t = (\sigma + \xi)\, \pi^h R^h_{kt} Q_{t-1} S_{t-1} - \sigma\pi^h R_t D_{t-1} \tag{3.32}$$

Depending on whether the constraint binds or not, aggregate net worth on each island will be or not relevant for general equilibrium dynamics.

For simplicity, the authors examine in more detail the solution for the two special cases of perfect ($\omega = 1$) and imperfect ($\omega = 0$) interbank markets.

Frictionless wholesale financial market ($\omega = 1$) The three optimality conditions (3.27)-(3.29) become:

$$(\upsilon_{bt} - \upsilon_t)\left(1 + \overline{\lambda}_t\right) = \theta\overline{\lambda}_t \tag{3.33}$$

$$\left(\frac{\upsilon_{st}}{Q^h_t} - \upsilon_{bt}\right)\left(1 + \lambda^h_t\right) = 0 \quad \Rightarrow \quad \frac{\upsilon_{st}}{Q^i_t} = \frac{\upsilon_{st}}{Q^n_t} = \upsilon_{bt} \tag{3.34}$$

[18] Equivalently, the ratio of total asset values to net worth are equalized $\left(\frac{Q^i_t S^i_t}{N^i_t} = \frac{Q^n_t S^n_t}{N^n_t}\right)$ and net interbank lending on in each island becomes zero.

$$
\begin{aligned}
\upsilon_t n_t^h &\geq \left[\theta - \left(\frac{\upsilon_{st}}{Q_t^h} - \upsilon_t\right)\right]\left(Q_t^h s_t^h - b_t^h\right) \\
&\geq (\theta - \mu_t)\left(Q_t^h s_t^h - b_t^h\right)
\end{aligned}
$$

$$
\phi_t n_t^h \geq Q_t^h s_t^h - b_t^h \tag{3.35}
$$

where $\mu_t = \frac{\upsilon_{st}}{Q_t^h} - \upsilon_{bt}$ is the excess marginal value of assets over deposits and $\phi_t = \frac{\upsilon_t}{\theta - \mu_t}$ is the maximum possible leverage ratio for asset purchases financed by deposit debt. The lower ϕ_t, the tighter the incentive constraint binds[19]. Condition (3.34) says that with a perfectly efficient interbank market, the shadow cost of interbank loans will be equated to the marginal value of the assets they are financing. Moreover, arbitrage in the wholesale market equalizes asset prices in both types of island, yielding symmetric results in both states[20]. For a sufficiently low level of net worth, condition (3.35) binds with equality and determines completely banks' demand for deposits, which is the same across islands.

Step 3 We can now solve for banks' shadow prices of financial assets in terms of the market rates of interest rates (which are determined in the general equilibrium). The Bellman equation (3.23) may now be rewritten as:

$$
V_t(s_t^h, b_t^h, d_t) = \max E_t \Lambda_{t,t+1} \sum_{h'=i,n} \pi^{h'} \left\{(1-\sigma) n_{t+1} + \sigma V_{t+1}(s_{t+1}^{h'}, b_{t+1}^{h'}, d_{t+1})\right\}
$$

where $n_{t+1} = R_{kt+1} Q_t s_t - R_{bt+1} b_t - R_{t+1} d_t$. The shadow prices of assets today determine next period's net worth, which in turn determines how tight the constraint will bind in the next period. Thus to evaluate the marginal values of assets today, we need to find the shadow value of net worth in the next period. Substituting the flow-of-fund constraint (3.20) and the optimality conditions (3.34) and (3.35) into the Bellman equation such that:

$$
\begin{aligned}
V_{t+1}(s_{t+1}^{h'}, b_{t+1}^{h'}, d_{t+1}) &= \upsilon_{st+1} s_{t+1}^{h'} - \upsilon_{bt+1} b_{t+1}^{h'} - \upsilon_{t+1} d_{t+1} \\
&= \upsilon_{st+1}\left(\frac{\phi_{t+1} n_{t+1} + b_{t+1}^{h'}}{Q_{t+1}}\right) - \upsilon_{bt+1} b_{t+1}^{h'} - \upsilon_{t+1}\left[\left(\phi_{t+1} n_{t+1} + b_{t+1}^{h'}\right) - n_{t+1} - b_{t+1}^{h'}\right] \\
&= \left(\mu_{t+1}\phi_{t+1} + \upsilon_{t+1}\right) n_{t+1}
\end{aligned}
$$

The marginal value of banks' net worth in period t+1 is equal to:

$$
\Omega_{t+1} = 1 - \sigma + \sigma\left(\upsilon_{t+1} + \mu_{t+1}\phi_{t+1}\right) \tag{3.36}
$$

which is a weighted average of the marginal values for exiting and surviving banks in the next period. In the latter case, the expression in bracket describes the net gain from having an additional unit of net worth in the next period: the first term (υ_{t+1}) represents the marginal cost of deposits and the second term ($\mu_{t+1}\phi_{t+1}$) represents the marginal gain from relaxing the

[19] As in the previous section (general case), the incentive constraint is tighter, the larger θ and the lower μ_t.

[20] This simplifies the analysis as it means that the returns on assets are equal, and therefore bankers' net worth across islands.

incentive constraint.

Then, assuming the guess of the (linear) value function (3.25) is correct, the following equality must hold:

$$\upsilon_{st} s_t^h - \upsilon_{bt} b_t^h - \upsilon_t d_t = E_t \Lambda_{t,t+1} \Omega_{t+1} \left(R_{kt+1} Q_t s_t - R_{bt+1} b_t - R_{t+1} d_t \right)$$

Using the method of undetermined coefficients we find that:

$$\upsilon_t = E_t \Lambda_{t,t+1} \Omega_{t+1} R_{t+1} \tag{3.37}$$

$$\mu_t = E_t \Lambda_{t,t+1} \Omega_{t+1} \left(R_{kt+1} - R_{t+1} \right) \tag{3.38}$$

$$E_t \Lambda_{t,t+1} \Omega_{t+1} R_{kt+1} = E_t \Lambda_{t,t+1} \Omega_{t+1} R_{bt+1} \tag{3.39}$$

where $\Lambda_{t,t+1}\Omega_{t+1}$ is an augmented stochastic discount factor[21].

To determine the aggregate implications of the presence of financial frictions in the general equilibrium, we need to derive the aggregate demand for bank assets as well as an expression for aggregate net worth.

First, we note that since condition (3.35) applies to all banks in the economy and ϕ_t is independent of bank specific factors, we can aggregate this condition over all banks to obtain:

$$Q_t S_t = \phi_t N_t \tag{3.40}$$

It says that aggregate bank capital affects the aggregate demand for bank assets, or equivalently the aggregate demand for capital goods (and therefore investment), in a proportion that depends on the excess value of bank assets (firms' securities) and the parameter θ.

Finally, since in this case banks in both islands are capital constrained and asset prices are identical so that net worth (per unit of asset) is identical across islands, the relevant state variable is the aggregate net worth of banks in the whole economy:

$$N_t = \left(\sigma + \xi \right) R_{kt} Q_{t-1} S_{t-1} - \sigma R_t D_{t-1} \tag{3.41}$$

To summarize the aggregate demand for bank assets which govern aggregate investment jointly with the capital-supply schedule[22]:

$$Q_t S_t = \phi_t N_t$$

$$\phi_t = \frac{\upsilon_t}{\theta - \mu_t}$$

$$\upsilon_t = E_t \Lambda_{t,t+1} \Omega_{t+1} R_{t+1}$$

[21] Substituting the expression for Ω_{t+1} into (3.37), we obtain the following expression: $\upsilon_t = E_t \Lambda_{t,t+1} \left[1 - \sigma + \sigma \left(\mu_{t+1} \phi_{t+1} + \upsilon_{t+1} \right) \right] R_{t+1}$. Thus marginal (shadow) values are forward-looking(?) parameters.

[22] Need to add the FOC for capital-producing firms to determine the price and (new) assets.

$$\mu_t = E_t \Lambda_{t,t+1} \Omega_{t+1} \left(R_{kt+1} - R_{t+1} \right)$$

$$\Omega_{t+1} = 1 - \sigma + \sigma \left(\mu_{t+1} \phi_{t+1} + \upsilon_{t+1} \right)$$

$$N_t = (\sigma + \xi) R_{kt} Q_{t-1} S_{t-1} - \sigma R_t D_{t-1}$$

$$D_{t-1} = Q_{t-1} S_{t-1} - N_{t-1}$$

$$R_{kt+1} = \psi_{t+1} \frac{\alpha Y_{t+1}/K_{t+1} + (1-\delta) Q_{t+1}}{Q_t}$$

In addition, two market clearing conditions are directly related to the problem examined in this section, one for firms' securities (S_t) and one for bank deposits (D_t)[23]:

$$S_t = K_t \tag{3.42}$$

$$D_{ht} = D_t \tag{3.43}$$

where D_{ht} is households' demand for deposits.

Symmetric frictions in the retail and wholesale financial markets ($\omega = 0$) The three optimality conditions (3.27)-(3.29) become:

$$\left(\upsilon_{bt} - \upsilon_t \right) \left(1 + \overline{\lambda}_t \right) = 0 \quad \Rightarrow \quad \upsilon_{bt} = \upsilon_t \tag{3.44}$$

$$\left(\frac{\upsilon_{st}}{Q_t^h} - \upsilon_{bt} \right) \left(1 + \lambda_t^h \right) = \theta \lambda_t^h \quad \Rightarrow \quad \mu_t^i > \mu_t^n \geq 0 \quad \Leftrightarrow \quad Q_t^i < Q_t^n \tag{3.45}$$

$$\upsilon_t n_t^h \geq \left[\theta - \left(\frac{\upsilon_{st}}{Q_t^h} - \upsilon_t \right) \right] Q_t^h s_t^h$$

$$\phi_t^h n_t^h \geq Q_t^h s_t^h \tag{3.46}$$

Condition (3.44) says that if deposits and interbank loans are perfect substitutes for banks, arbitrage should equalize their marginal costs in equilibrium. However, as described by condition (3.45), this case poses the most severe limits to arbitrage in the interbank market, associated with the highest gap between excess values across islands, or equivalently the highest price dispersion. Here we assume that the incentive constraint tights more strictly in investing islands so that the excess value on these islands should be larger[24]. Condition (3.46) says that, in this case, the incentive constraint applies to the entire stock of bank assets (instead of those financed by deposits in the previous case). In addition, banks on non-investing islands may or may not be constrained, depending on whether or not they are able to reduce the excess value to zero through arbitrage between securities and interbank loans. If they are able to do so, then banks

[23] To close the RBC model a labour market clearing condition is also required.

[24] This is because there are more securities per unit of net worth in investing islands and, due to imperfections in the interbank market, these additional securities can only be funded if they promise a lower return than securities in the interbank market, leading to an inefficient allocation of resources between the two types of islands.

on investing islands will be able to borrow some funds from banks in non-investing islands; otherwise, there is no interbank lending in equilibrium[25]. Thus condition (3.46) implies:

$$Q_t^i s_t^i = \phi_t^i n_t^i \tag{3.47}$$

where $\phi_t^i = \frac{\upsilon_t}{\theta - \mu_t^i}$, and

$$Q_t^n s_t^n \leq \phi_t^n n_t^n, \qquad \left(Q_t^n s_t^n - \phi_t^n n_t^n\right) \mu_t^n = 0 \tag{3.48}$$

where $\phi_t^n = \frac{\upsilon_t}{\theta - \mu_t^n}$.

Step 3 Solving for banks' shadow prices of financial assets in terms of the market rates of interest rates. The Bellman equation (3.23) may now be rewritten as:

$$V_t(s_t^h, b_t^h, d_t) = \max E_t \Lambda_{t,t+1} \sum_{h'=i,n} \pi^{h'} \left\{(1-\sigma) n_{t+1}^{h'} + \sigma V_{t+1}(s_{t+1}^{h'}, b_{t+1}^{h'}, d_{t+1})\right\}$$

where $n_{t+1}^{h'} = R_{kt+1}^{hh'} Q_t^h s_t - R_{bt+1} b_t - R_{t+1} d_t$ (to revise: how do we account for the fact that net worth is redistributed at the beginning of periof t+1 to equalize the ex-ante expected return to bank assets, $R_{kt+1}^{hh'}$?). The only difference with the previous case is that now, because asset prices are state-dependent, next period net worth depends on the future state in each island so that the result are conditional on the specific realization of the liquidity shock (h'). The solution method is the same as previously and gives:

$$\Omega_{t+1}^{h'} = 1 - \sigma + \sigma \left(\upsilon_{t+1} + \mu_{t+1}^{h'} \phi_{t+1}^{h'}\right) \tag{3.49}$$

$$\upsilon_t = E_t \Lambda_{t,t+1} \sum_{h'=i,n} \pi^{h'} \Omega_{t+1}^{h'} R_{t+1} \tag{3.50}$$

$$\mu_t^h = E_t \Lambda_{t,t+1} \sum_{h'=i,n} \pi^{h'} \Omega_{t+1}^{h'} \left(R_{kt+1}^{hh'} - R_{t+1}\right) \tag{3.51}$$

$$\begin{aligned} E_t \Lambda_{t,t+1} \sum_{h'=i,n} \pi^{h'} \Omega_{t+1}^{h'} R_{kt+1}^{ih'} &> E_t \Lambda_{t,t+1} \sum_{h'=i,n} \pi^{h'} \Omega_{t+1}^{h'} R_{kt+1}^{nh'} \\ &\geq E_t \Lambda_{t,t+1} \sum_{h'=i,n} \pi^{h'} \Omega_{t+1}^{h'} R_{bt+1} = E_t \Lambda_{t,t+1} \sum_{h'=i,n} \pi^{h'} \Omega_{t+1}^{h'} R_{t+1} \end{aligned} \tag{3.52}$$

where the second inequality holds with equality if $\mu_t^n = 0$.

The aggregate demand for bank assets in the two types of islands are obtain by aggregating conditions (3.47) and (3.54) over all type i and type n islands

$$Q_t^i S_t^i = \phi_t^i N_t^i \tag{3.53}$$

[25] Which parameters affect this result: the steady state spread between returns on securities and deposit rate, the steady state leverage ratio,...?

$$Q_t^n S_t^n \leq \phi_t^n N_t^n, \qquad \left(Q_t^n S_t^n - \phi_t^n N_t^n\right)\mu_t^n = 0 \tag{3.54}$$

Imperfections in the interbank market amplify the extent to which banks in islands i are financially constrained through their effect on asset prices: lower asset prices in islands i lead to a more than one-for-one fall in banks' net worth (because they are leveraged), which pushes asset prices further down, and so on. Aggregate bank net worth in each island is given by (3.32). To study the implications of credit market frictions for the real economy (investment) we need only take care of the balance sheets of banks on investing islands since it is asset prices on these islands that determine the price of new capital goods, and therefore the level of investment.

To summarize the aggregate demand for bank assets which govern aggregate investment jointly with the capital-supply schedule[26]:

$$Q_t^i S_t^i = \phi_t^i N_t^i$$

$$\phi_t^i = \frac{\upsilon_t}{\theta - \mu_t^i}$$

$$\upsilon_t = E_t \Lambda_{t,t+1} \sum_{h'=i,n} \Omega_{t+1}^{h'} R_{t+1}$$

$$\mu_t^i = E_t \Lambda_{t,t+1} \sum_{h'=i,n} \pi^{h'} \Omega_{t+1}^{h'} \left(R_{kt+1}^{ih'} - R_{t+1}\right)$$

$$\Omega_{t+1}^{h'} = 1 - \sigma + \sigma\left(\upsilon_{t+1} + \mu_{t+1}^{h'} \phi_{t+1}^{h'}\right)$$

$$N_t^i = \pi^i\left[(\sigma + \xi) R_{kt}^i Q_{t-1} S_{t-1} - \sigma R_t D_{t-1}\right]$$

$$D_{t-1} = Q_{t-1} S_{t-1} - N_{t-1}$$

$$R_{kt}^i = \psi_t \frac{\alpha Y_t / K_t + (1-\delta) Q_t^i}{Q_{t-1}}$$

In addition, two market clearing conditions are directly related to the problem examined in this section, one for firms' securities (S_t) and one for bank deposits (D_t):

$$S_t^i = I_t + (1-\delta)\pi^i K_t \tag{3.55}$$

$$D_{ht} = D_t \tag{3.56}$$

3.2.3 The propagation of shocks in the model economy

Gertler and Kiyotaki (2011) investigate how frictions in the retail and wholesale markets may have real effects on key macroeconomic variables and through which channels. They propose a 'crisis' experiment whereby an exogenous, proacted, capital quality shock (ψ_t) triggers a decline in banks'asset returns in period t that is mutually reinforced by two mechanisms. First, the decline in banks' net worth tightens banks' incentive constraint and leads to a decline in net

[26]Need to add the FOC for capital-producing firms to determine the price and new assets.

worth per unit of assets, which pushes asset prices down ('fire sale of assets'). Second, the fall in asset prices is transmitted more than one for one to net worth (because of leverage), thus further tightening the incentive constraint. The fall in asset prices corresponds to an increase in the excess value of bank assets necessary for banks to obtain funds and leads to an increase in the spread between loan rates and deposit rates. Since the cost of financing new capital purchases increases for non-financial firms, their demand for new capital goods falls. This explains the significant fall in investment and the much stronger effect of the shock on output in the models with frictions compared to the frictionless RBC model. The former models also exhibit more persistence, which is driven by the dynamic of net worth. For the spreads to fall back to their steady state values, banks must build up their net worth per unit of assets again. As can be seen from the definition of ϕ_t, banks' leverage is procyclical so that they must be deleveraging during the recovery. In the case of an imperfect interbank market, these effects are amplified because the incentive constraint binds more strongly in investing islands, thus leading to a larger fall in asset prices and net worth and a higher increase in bank leverage on impact.

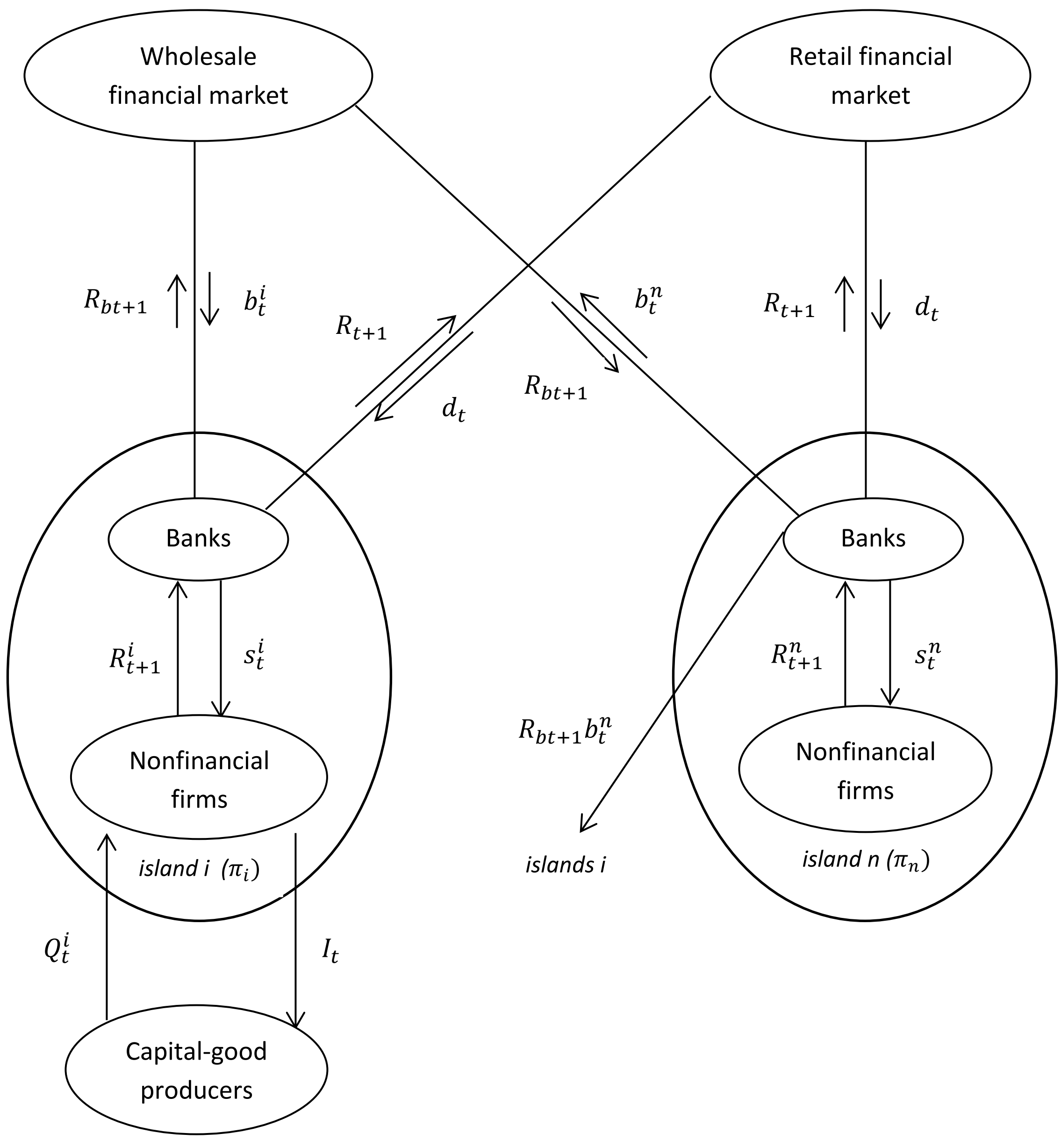


Figure 3-3: Financial intermediation in Gertler and Kiyotaki (2011)

Chapter 4

An investigation into the determinants of the financial accelerator effect in BGG

Abstract *Bernanke, Gertler and Gilchrist (1999) developed a benchmark model to study how a microfunded financial friction which depends on borrowers' balance sheet conditions may account for the behaviour of key macroeconomic variables via its effect on investment. They show how a financial accelerator effect, characterized both by amplification and persistence, may emerge and be quantitatively significant. This mechanism is embedded into a New Keynesian framework in order to provide a credible account of the transmission of a monetary shock. Their work fit into a research program which saw the development of competing quantitative models of financial frictions (Carlstrom and Fuerst, 1997; Kiyotaki and Moore, 1997) and was regenerated by the recent financial crisis. This study proposes to investigate a neglected aspect of the intuition provided by BGG, the interaction between the slope of the Philips curve and the financial accelerator effect. We show that the magnitude of the financial accelerator effect in BGG depends crucially on the value of the Calvo parameter. In particular, the greater this value, the flatter the Philips curve and the stronger is the financial accelerator effect.*

4.1 Introduction

The framework is the neoclassical growth model featuring households and firms. The economic environment is modified so that the following set of questions can be addressed. How important are fluctuations in borrowers financial conditions for business cycle fluctuations? How does the financial accelerator effect interact with the behaviour of asset prices? How are monetary and fiscal shocks transmitted in such a model economy?

The environment is described in Figure 4-4. There are three types of firms. Competitive entrepreneurs[1] produce the wholesale good and are affected by a funding constraint in their decision to purchase capital assets. Wholesale goods are sold to monopolistic retailers whose ability to change prices is time-dependent. This New Keynesian feature allows studying the effect of a monetary shock. Capital goods are produced and sold by competitive capital-producing firms facing adjustment costs. This feature allows for an amplification effect through fluctuations in asset prices.

Lenders channel funds between households and entrepreneurs by taking deposits (which constitute the only form of savings for households) and issuing loans. Finally, households make labour and consumption decisions. The equilibrium is characterized by two labour market clearing conditions as well as a retail goods market clearing condition and a debt (deposit) market clearing condition which says that total saving by households equals total borrowing by entrepreneurs.

To complete the description of the model, we specify the functional forms describing households' preferences, entrepreneurs' and capital-producing firms' production functions and their associated parameters which are consistent with balanced growth. Retailers' pricing behaviour is assumed to follow a Calvo (1983) model in which the Calvo parameter determines the frequency at which retailers are allowed to change their prices. The aggregate external finance schedule defined as the outcome of the optimal financial contract gives rise to an additional optimality condition. Its derivation makes use of the linearity of the monitoring technology. Finally, the model is completed by a monetary rule defined as a standard Taylor rule.

Next, the model is linearized around its steady state along the lines exposited by Uhlig (1995). In doing so, we restrict the behaviour of some key variables determining the optimal financial contract to remain constant over the business cycle. Such a procedure also eschews questions related to the possibility of multiple equilibria. In particular, it would be ill-suited to study the effects of rare events. The key econometric problem resides in the calibration of the parameters pertaining to the financial conditions in the entrepreneurial sector. By carefully justifying how they can be defined and measured econometrically using actual data and by restricting their steady state outcome, this procedure makes it possible to study the behaviour of an economy characterized by such properties. Kydland and Prescott (1991) described this exercise as 'the econometrics of the general equilibrium approach'.

The linearized and calibrated model is solved numerically in Dynare (MATLAB) using

[1] Assuming entrepreneurs are competitive facilitates the aggregation of this constraint over the entrepreneurial sector as a whole.

the method developed by Blanchard and Khan (1980). This method looks for the saddle path of the system before solving for the recursive law of motion of the endogenous variables. Once this is done, we proceed with the computational experiment consisting in examining the behaviour of the endogenous variables as they respond to a monetary policy shock using impulse response analysis. This methodology allows us, providing the calibration process respected some elementary rules (Cooley, 1997), to compare the behaviour of different model economies following an exogenous policy shock.

In this chapter, we propose to examine the relationship between the slope of the Philips curve and the financial accelerator effect. The Philips curve in this model, states that current inflation is increasing in the current real marginal cost (the strength of this relationship depending on the Calvo parameter) and expected future inflation. The question we propose to answer takes the following form. In the BGG model, how sensitive is the financial accelerator effect to the magnitude of the Calvo parameter? How much of the financial accelerator effect can be explained by the interaction between the financial friction and inflation dynamics. One important implication of the presence of the financial friction is that it adds persistence to the system. One can therefore also ask how much of this persistence effect can be explained by the strength of the relationship between the slope of the Philips curve and the financial accelerator mechanism.

4.2 Equilibrium conditions

In this section, we present the framework in more detail, describe the model and the choices of functional forms and derive the optimal conditions and market-clearing conditions. Second, we clarify the grounds on which the interaction between the financial accelerator mechanism and the assumption of price stickiness can be investigated. On one side, the Calvo model establishes a relationship between the current real marginal cost and the inflation rate (Philips curve). On the other side, introducing a real financial friction makes the cost of external finance for entrepreneurs countercyclical, thus affecting the dynamics of the real marginal cost over the business cycle and potentially inflation. Is this framework well suited to think about these interactions? Can it be modified to study these interactions? To what extent the answer to this question depends on the way the labour market has been modelled?

4.2.1 Households

Households are infinitely lived. In each period they sell their labour to entrepreneurs who produce the wholesale good and choose how much to consume of the retail good and how much to deposit at the financial intermediaries (paying a real riskless rate of return). They are risk-averse and require the interest rate on their deposit to be risk-free. Their utility is separable in real consumption, real money balances and leisure and takes the following form:

$$U(C_t, M_t/P_t, H_t) = \ln C_t + \varsigma \ln(M_t/P_t) + x \ln(1 - H_t) \tag{4.1}$$

C_t is real consumption, M_t/P_t is real money balance and H_t is households' labour supply (the fraction of their time endowment spent working). Money is incorporated through a money-in-the-utility function; this feature allows the model to generate a liquidity effect in equilibrium[2].

Households' stream of income in each period is composed of their labour income, the return they receive from their deposits at financial intermediaries from the previous period, the dividends they receive from retail firms that they own, their real money balance holdings from the previous period, discounted by the current period inflation rate and finally government lump-sum taxes which are used to finance the money supply and government expenditures. The households' budget constraint then is:

$$C_t + D_t + \frac{M_t}{P_t} = W_t H_t - T_t + \Pi_t + (1 + r_{t-1}) D_{t-1} + \frac{M_{t-1}}{P_{t-1}} \left(\frac{1}{1 + \pi_t} \right) \tag{4.2}$$

π_t is the inflation rate, W_t is the real wage, D_t are real deposits, Π_t are dividends, T_t are lump sum taxes, r_{t-1} is the real risk-free interest rate.

Households solve for the infinite sequences of real consumption, real money balances, real deposits and labour supply:

$$\max_{C_t, M_t/P_t, H_t, D_t} E_t \sum_{k=0}^{\infty} \beta^k \left[\ln C_{t+k} + \varsigma \ln \left(\frac{M_{t+k}}{P_{t+k}} \right) + \xi \ln(1 - H_{t+k}) \right] \tag{4.3}$$

subject to (4.2).

The first-order conditions for real consumption and labour supply yield two optimality conditions, the consumption Euler equation and the labour supply schedule. Consumption Euler equation:

$$\frac{1}{C_t} = \beta R_t E_t \left(\frac{1}{C_{t+1}} \right) \tag{4.4}$$

Labour supply:

$$W_t = \xi \left(\frac{1}{1 - H_t} \right) \tag{4.5}$$

The optimal level of deposits is determined by the debt market-clearing condition.

4.2.2 Entrepreneurs

Entrepreneurs survive between period t-1 and period t with probability λ; their expected lifetime is therefore equal to $1/(1-\lambda)$. When entrepreneurs die, they are automatically replaced so that there is a constant fraction of entrepreneurs in the economy. Entrepreneurs are competitive and use a Cobb-Douglas production function to produce the wholesale good (CRS):

$$Y_t = A_t K_t^{\alpha} L_t^{1-\alpha} \tag{4.6}$$

[2]This specification of the utility function is a special case of the CES functional form where the intertemporal elasticity of substitution between real consumption and leisure is set equal to one.

K_t is the capital stock purchased by entrepreneurs in period t-1. Total labour is supplied both by households and entrepreneurs so that aggregate labour is defined as follows:

$$L_t = H_t^{\Omega}(H_t^e)^{1-\Omega} = H_t^{\Omega} \tag{4.7}$$

H_t^e is entrepreneurial labour. It is supplied inelastically and normalised to 1.

Entrepreneurs' demand schedules The expected return per unit of capital expenditure R_t^k is determined by the marginal product of capital and capital gains and is expressed in terms of the retail price. The demand curve for capital expenditures is given by:

$$E(R_{t+1}^k) = E\left[\frac{\frac{1}{X_{t+1}}\frac{\alpha Y_{t+1}}{K_{t+1}} + Q_{t+1}(1-\delta)}{Q_t}\right] \tag{4.8}$$

δ is the rate of depreciation of capital, X_t is the retailers markup so that $1/X_t$ is the relative price of wholesale goods and Q_t is the relative price of capital goods charged by capital-producing firms.

The demand schedules for household and entrepreneurial labour are:

$$W_t = \frac{\Omega(1-\alpha)Y_t}{X_t H_t} \quad \text{and} \quad W_t^e = \frac{(1-\Omega)(1-\alpha)Y_t}{X_t} \tag{4.9}$$

W_t and W_t^e are total wage income received by household and entrepreneurs. Households and entrepreneurs' real wages (relative to the price of the retail good) are equal to their respective marginal products (or contributions to the production of the wholesale good).

Entrepreneurs' preferences imply that they reinvest all of their net worth into their next period capital stock. However, those exiting at the end of period t consume their profits. Aggregate entrepreneurial consumption is:

$$C_t^e = (1-\lambda)V_t \tag{4.10}$$

Thus entrepreneurs as a whole consume a constant fraction of their profits[3].

The external finance schedule In chapter 2, we derived both the external finance schedule and the expression for aggregate net worth. The first describes how the financial friction is introduced in the entrepreneurs' problem:

$$E(R_{t+1}^k) = \chi\left(\frac{N_{t+1}}{Q_t K_{t+1}}\right) R_t \quad \text{with} \quad \chi'\left(\frac{N_{t+1}}{Q_t K_{t+1}}\right) < 0 \tag{4.11}$$

This external finance schedule determines the lowest value of the expected return to capital necessary to purchase a given amount of capital expenditure ($Q_t K_{t+1}$). It is inversely proportional

[3]The real wage of exiting entrepreneurs is not included in this expression. It is automatically transferred to new entrepreneurs whose net worth is just equal to this real wage.

to entrepreneurs' net worth. Aggregate net worth is a state variable:

$$N_{t+1} = \lambda[(1-\mu G(\overline{\omega}_t))R_t^k Q_{t-1}K_t - R_t(Q_{t-1}K_t - N_t)] + \frac{(1-\Omega)(1-\alpha)Y_t}{X_t} \tag{4.12}$$

4.2.3 Capital-producing firms

The problem faced by capital-producing firms was described in chapter 2. They are competitive and face adjustment costs, making the price of capital increase in the ratio of aggregate investment to capital. The price of capital (Tobins q) is given by:

$$Q_t = \frac{1}{\phi'(I_t/K_t)} \quad \text{and} \quad \frac{\partial Q_t}{\partial(I_t/K_t)} = -\frac{\phi''}{\phi'^2} > 0 \tag{4.13}$$

4.2.4 Retailers

Retailers are monopolists. They buy the wholesale good from entrepreneurs at the competitive price, differentiate it at no cost and resell it at a markup. Thus, the retail good is a composite good:

$$Y_t^f = \left[\int_0^1 Y_t(z)^{(\epsilon-1)/\epsilon}\,\mathrm{d}z\right]^{(\epsilon-1)/\epsilon} \tag{4.14}$$

Y_t^f represents the total quantity of retail good produced, $Y_t(z)$ is the quantity of output sold by retailer z and ϵ is the elasticity of demand with respect to retail goods. Households (and entrepreneurs) purchase the combination of individual retail goods that minimizes the total expenditure associated with the consumption of any amount of the composite good Y_t^f. The outcome of this problem is given by the demand curve for retail goods:

$$Y_t(z) = \left(\frac{P_t(z)}{P_t}\right)^{-\epsilon} Y_t^f \tag{4.15}$$

$P_t(z)$ is the price of individual retail goods and P_t is the aggregate price level. The aggregate demand for retail goods is proportional to the aggregate demand for the composite good and depends on their relative price and the elasticity of demand.

Optimal pricing We assume some degree of price stickiness and introduce this idea by using the Calvo (1983) model. According to this approach, there is a fraction θ of retailers who keep their prices unchanged in each period instead of setting the new profit-maximizing price P_t^*. The new optimal price in each period maximizes profits for the entire duration it is expected to remain unchanged, subject to the set of demand curves associated with this price in each of these periods. The solution to this problem is given by:

$$\sum_{k=0}^{\infty} \theta^k E_t\left[\frac{1}{R_{t+k}} Y_{t+k}^* \left(P_t^* - \frac{X}{X_{t+k}} P_{t+k}\right)\right] = 0 \tag{4.16}$$

Y^*_{t+k} is the output given the price P^*_t, $X = \epsilon/(\epsilon - 1)$ is the steady-state markup, $1/X_t$ is the real marginal cost which is equal to the inverse of the markup (it is the price charged by entrepreneurs for wholesale goods). This condition says that the new profit-maximizing price is such that in expectation, the discounted marginal revenue equals the sum of discounted real marginal costs in each period the new price is expected to remain unchanged.

4.2.5 Market-clearing conditions

There are four markets, two labour markets, a retail good market and a debt market. Labour markets are competitive and in equilibrium demand equals supply. Thus, Equilibrium households' employment is determined jointly by the demand schedule for labour (LHS) and the schedule for household labour supply (RHS):

$$\frac{\Omega(1-\alpha)Y_t}{X_t H_t} = \xi \left(\frac{1}{1-H_t}\right) C_t \tag{4.17}$$

Since entrepreneurs' labour supply is inelastic, the entrepreneurs' labour market clearing condition says that their real wage fluctuates to accommodate variations in entrepreneurs' demand for labour.

The goods market equilibrium is given by the aggregate resource constraint:

$$Y^f_t = C_t + C^e_t + I_t + G_t + \mu \int_0^{\overline{\omega}_t} \omega \, \mathrm{d}F(\omega) R^k_t Q_{t-1} K_t \tag{4.18}$$

$I_t = K_{t+1} - (1-\delta)K_t$ is aggregate investment and $\mu \int_0^{\overline{\omega}_t} \omega \, \mathrm{d}F(\omega) R^k_t Q_{t-1} K_t$ is the deadweight loss generated by the financial friction. Equilibrium in the market for loanable funds requires that the total deposits held by households be equal to the total amount borrowed by entrepreneurs:

$$D_t = Q_t K_{t+1} - N_{t+1} \tag{4.19}$$

4.3 Loglinearization

In the set of log-linearized equations, percentage deviations from the steady state are denoted by small latin letters.

Aggregate demand

Goods market equilibrium:

$$y_t = \left(\frac{C}{Y}\right) c_t + \left(\frac{I}{Y}\right) i_t + \left(\frac{G}{Y}\right) g_t + \left(\frac{C^e}{Y}\right) c^e_t + \frac{DR^k K}{Y}(r^k_t + q_{t-1} + k_t) \tag{4.20}$$

$$D = 1 - \mu \int_0^{\overline{\omega}_t} \omega \, \mathrm{d}F(\omega) \tag{4.21}$$

Consumption Euler equation:

$$c_t = -r_t + E_t(c_{t+1}) \tag{4.22}$$

Entrepreneurs consumption:

$$c_t^e = n_{t+1} \tag{4.23}$$

External finance premium:

$$E_t(r_{t+1}^k) - r_t = -\chi[n_{t+1} - (q_t + k_{t+1})] \tag{4.24}$$

$$\chi = \frac{\Psi(R^k/R)}{\Psi'(R^k/R)}\frac{R}{R^k} \tag{4.25}$$

The parameter χ is evaluated at the steady state. It is a function of the steady state cutoff value of entrepreneurs' idiosyncratic shock $\overline{\omega}^{ss}$.
Return to capital:

$$r_{t+1}^k = (1-\varepsilon)[y_t - k_{t+1} - x_{t+1}] + \varepsilon q_{t+1} - q_t \tag{4.26}$$

Tobin's q:

$$q_t = \varphi(i_t - k_{t+1}) \tag{4.27}$$

Aggregate supply
Production function:

$$y_t = a_t + \alpha k_t + (1-\alpha)\Omega h_t \tag{4.28}$$

Labour supply:

$$y_t - h_t - x_t - c_t = \eta^{-1} h_t \tag{4.29}$$

Phillips curve:

$$\pi_t = \kappa(-x_t) + \beta E_t(\pi_{t+1}) \tag{4.30}$$

$$\kappa = \left(\frac{1-\theta}{\theta}\right)(1-\theta\beta) \tag{4.31}$$

Evolution of state variables

Capital accumulation:

$$k_{t+1} = \delta i_t + (1-\delta)k_t \tag{4.32}$$

Entrepreneurs' net worth:

$$n_{t+1} = \left(\frac{\lambda RK}{N}\right)(r_t^k - r_{t-1}) + r_{t-1} + n_t \tag{4.33}$$

This is the expression used in BGG and this preliminary analysis. We propose to use the more accurate expression in future work:

$$\begin{aligned} n_{t+1} = &\left(\lambda DR^k \frac{K}{N}\right) r_t^k + \frac{\lambda}{\beta}\left(1 - \frac{K}{N}\right) r_{t-1} + \left(\frac{\lambda}{\beta}\right) n_t \\ &+ \left(\lambda \frac{K}{N}\right)\left(DR^k - \frac{1}{\beta}\right)(q_{t-1} - k_t) + \frac{(1-\alpha)(1-\Omega)}{\alpha}\frac{K}{XN}(y_t - x_t) \end{aligned} \tag{4.34}$$

Monetary policy rule and shock processes
Policy rule:

$$r_t^n = \rho r_{t-1}^n + \zeta \pi_{t-1} + \epsilon_t^{rn} \tag{4.35}$$

Fisher equation:

$$r_t = r_t^n - E_t(\pi_{t+1}) \tag{4.36}$$

Technology shock processes:

$$a_t = \rho_a a_{t-1} + \epsilon_t^a \tag{4.37}$$

$$\tag{4.38}$$

4.4 Parameter choices

The calibration exercise consists in matching the parameters of the model so as to ensure that the model is consistent with well documented facts about the properties of actual economies (here the US economy). Steady state outcomes should be consistent with the stylized facts characterizing a balanced growth path as evidenced by Solow (1970). Other parameters are set according to the literature, based micro-evidence documenting the behaviour of participants in the labour market (labour supply elasticity), the constraints faced investment firms in the production process (adjustment cost parameter) and the pricing behaviour of retail firms (Calvo parameter). Parameters in the monetary policy rules should be set to take into account both the existing theory modelling policy-makers behaviour and the econometric evidence documenting the associated stochastic processes for monetary variables. Finally, the originality of this model resides in the presence of parameters driving the relationship between borrower-entrepreneurs and lenders.

The parametrization strategy adopted consists therefore in first pinning down those last critical parameters. The values of the deep parametersγ, μ and σ, representing respectively the survival rate of entrepreneurs, default costs and the volatility of the idiosyncratic shock are set to imply three key steady state outcomes. First, the spread between entrepreneurs internal return to net worth and the risk-free rate is set to 200 basis points. Second, the steady state cutoff value for the idiosyncratic shock is fixed such that the annualized default rate $F(\overline{\omega})$ is equal to 3%. Third, the steady state leverage ratio N/QK is equal to 0.5; in other words, in steady state, entrepreneurs are assumed to borrow half of the amount necessary to finance their capital expenditure.

The remaining parameters are set to satisfy the steady state relationships described below.

The steady state

$$R = \frac{1}{\beta}$$

$$Y^f = C + C^e + I + G + \mu \int_0^{\overline{\omega}_t} \omega \, \mathrm{d}F(\omega) R^k K$$

$$Q = 1$$

$$Y = \frac{AK^\alpha (H^\Omega)^{1-\alpha}}{X}$$

$$R^k = \chi \left(\frac{N}{K}\right) R$$

$$C^e = N$$

$$R^k = \frac{1}{X}\frac{\alpha Y}{K} + 1 - \delta$$

$$\frac{\Omega(1-\alpha)Y}{XH} = \xi \left(\frac{1}{1-H}\right) C$$

The full set of parameter choices is presented in table 4.1.

4.5 The transmission mechanism of monetary shocks in BGG

We begin by asking how a monetary shock is transmitted to the real economy in the baseline model. This gives us the opportunity to show how, in this model, business cycle fluctuations are amplified by the financial accelerator mechanism. We simulate the response of the model economy to a 25 basis point fall in the nominal interest rate. We also generate the response to the same shock in an economy characterized by the same steady state outcome but without the financial accelerator mechanism. In this case, the spread between the internal rate of return and the risk-free rate plays no part in the dynamics of the system.

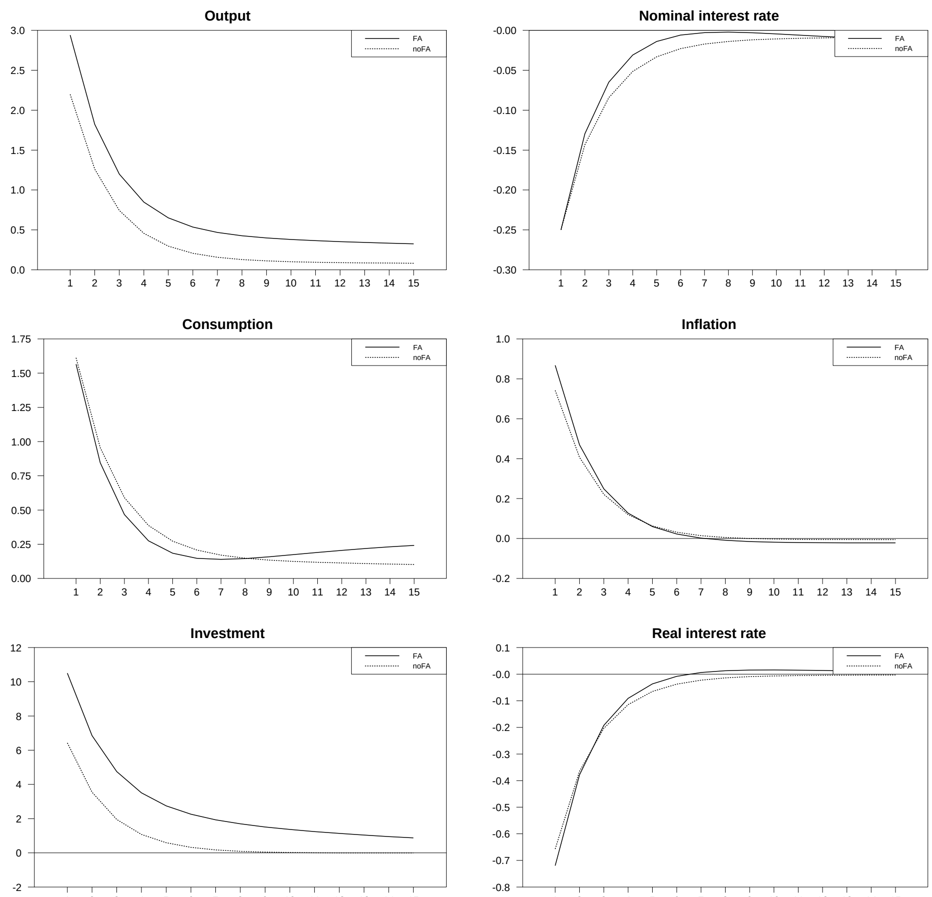
Figures 4.5 and 4.5 present the impulse response for several macroeconomic variables. Im-

	parameters	baseline specification
Optimal financial contract		
risk spread (ss)	$R^k - R$	0.0200
leverage (ss)	QK/N	2.08175
business failure rate	$F(\overline{\omega})$	0.03
survival probability	γ	0.9728
default cost	μ	0.12
variance of $\log \omega$	σ	0.28
cut-off rate for default	$\overline{\omega}$	0.490
Elasticity of EFP	χ	0.05209
Utility function		
discount rate	β	0.99
labour supply elasticity	$\eta = \frac{l}{H}$	3
Production function		
capital share	α	0.35
entrepreneurs' labour share of income	$(1-\alpha)(1-\Omega)$	0.01
households' labour share of income	$(1-\alpha)\Omega$	0.64
depreciation rate	δ	0.025
adjustment cost elasticity	$\varphi = -\phi''(I/K)I/K$	0.25
Retailers		
Calvo parameter	θ	0.75
Phillips curve parameter	κ	0.85833
markup (ss)	$X = \frac{\epsilon}{\epsilon - 1}$	1.1
Monetary policy		
autoregressive parameter	ρ	0.9
coefficient on inflation	ζ	0.11

Table 4.1: Model parameters

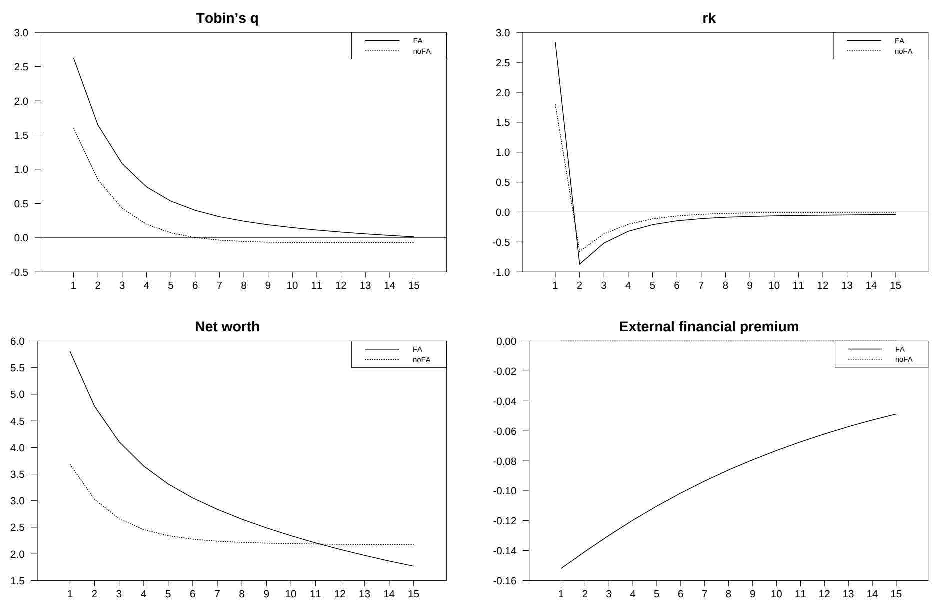
pulse responses are expressed in percentage deviations from the steady state, except for the financial variables which are expressed in basis point deviations. The impulse responses show how two main effects contribute to the transmission of the shock in the model economy. First, traditional intertemporal substitution effects determine the evolution of consumption and investment. This mechanism is governed by the respective Euler equations for consumption and investment. The second effect at work is the balance sheet channel. It states that the increase in investment boosts the price of capital (Tobin's q) and raises entrepreneurs net worth, thus reducing the external finance premium and further stimulating investment expenditure.

Figure 4-1: The financial accelerator effect



Impulse responses to a 25 basis point expansionary monetary shock. Results are shown for the model including a financial accelerator mechanism (FA) and a model without (noFA), characterized by a constant external finance premium (χ=0).

Figure 4-2: The financial accelerator effect (continued)



Impulse responses to a 25 basis point expansionary monetary shock. Results are shown for the model including a financial accelerator mechanism (FA) and a model without (noFA), characterized by a constant external finance premium (χ=0).

To see how these two effects interact with each other, it is necessary to analyse the behaviour of this economy in more detail. First, the persistent fall in the nominal interest rate is associated with a persistent increase in the inflation rate. As a result the real interest rate falls by more than the nominal interest rate. This leads households to substitute leisure away for labour and consumption increases (the same mechanism stimulates investment). Thus, in the model without financial accelerator, the dynamics of consumption, output and investment closely follow the path of the real interest rate. This result follows from the assumption of price stickiness and the Phillips curve relationship it is giving rise to (New Keynesian feature). Second, on impact, the rise in investment increases Tobin's q, which leads to an increase in the return to capital (r_k) and as a result to entrepreneurs' net worth. The external finance premium is inversely related to net worth and decreases. This reduces the cost of borrowing and boosts further investment. We note that because of the anticipated decline in Tobin's q, the growth of the return to capital becomes negative in the second period. However, since it declines by more than the change in the real interest rate, the external finance premium remains persistently lower.

4.6 The interaction between the financial accelerator effect and the slope of the Philips curve

To investigate whether the magnitude of the effects described in the previous section depend on the degree of price stickiness captured by the Calvo parameter θ, we systematically compare the strength of the financial effect for three different values of this parameter. In the baseline model, the value of this parameter was set to 0.75, which says that retailers change their prices on average once every three quarters. We want to know whether the results significantly change for relatively small variations in this parameter. These changes translate into different values for the slope of the Philips curve (κ). Table 4.2 describes these parameter changes. Since θ and κ are related in a nonlinear fashion (Table 4.1), small changes in the first leads to large swings in the second. The higher the degree of price stickiness (the larger θ), the flatter the Philips curve.

θ	κ
0.65	0.191961
0.75	0.085833
0.85	0.027970

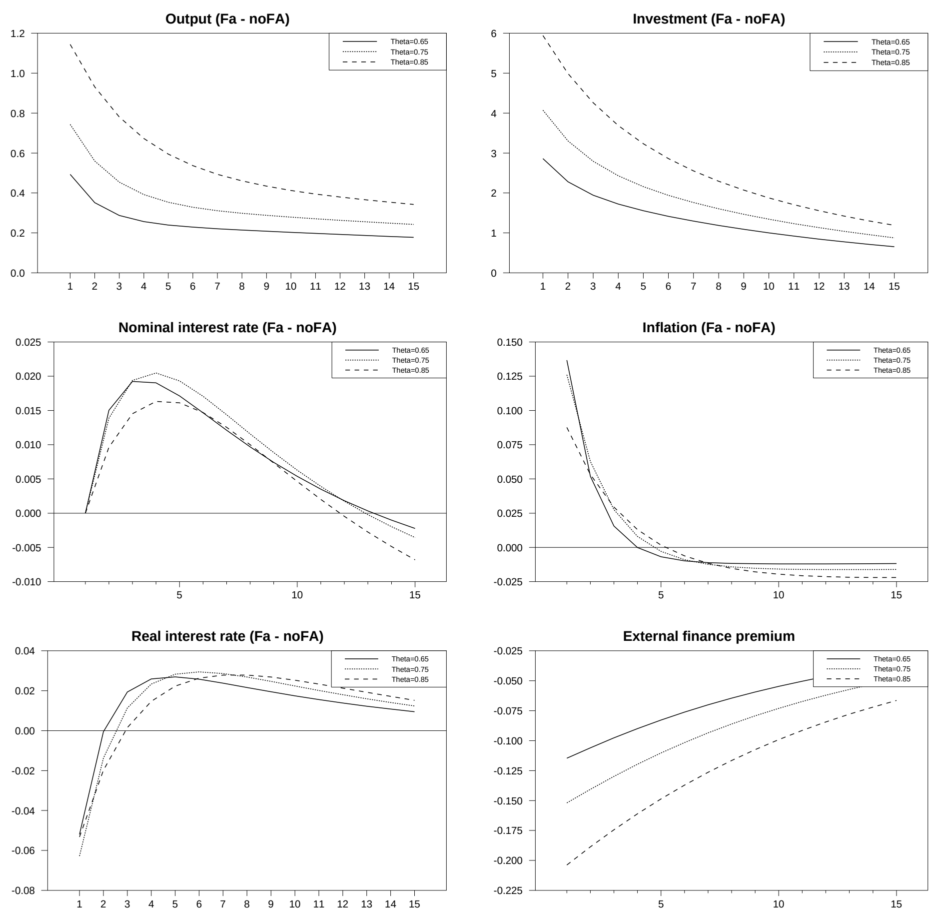
Table 4.2: The relationship between the degree of price stickiness and the slope of the Philips curve

Next, we generate three pairs of impulse responses to a 25 basis point fall in the nominal interest rate for each values of the Calvo parameter and for both the model with financial accelerator and the model without financial accelerator. For each pair, we then compute the difference between the two impulse responses (FA-noFA), so that only remains the deviation from the steady state that is due to the financial accelerator mechanism. For each value of θ, we therefore obtain a measure of the strength of the financial accelerator effect. We repeat the procedure four times, for output, investment, the nominal interest rate, inflation, the real interest rate and the external finance premium.

Figure 4.6 presents the results. It shows that for higher values of the Calvo parameter, the variation in output and investment that is due to the financial accelerator effect is significantly higher. In particular, by increasing its value from 0.75 to 0.85, the strength of this effect increases by about 50%. Similarly, the external finance premium decreases by more for higher values of the Calvo parameter. There is a 5 basis point difference on impact between a value of θ equal to 0.75 and a value of 0.85 (for an initial fall of the nominal interest rate of 25 basis points), which represents a 25% increase. In addition, for all three variables these differences are persistent.

Generally, a higher value of the Calvo parameter leads to smaller inflation on impact as it takes more time for prices to adjust. In turn, this leads to a weaker response of the nominal interest rate which remains lower throughout. However this mechanism is harder to observe for lower values of theta as the cases with theta equal to 0.65 and theta equal to 0.75. Inspection of the behaviour of the real interest may provide more insight into the effect examined in this experiment. The change in theta does not affect the magnitude of the financial accelerator effect on the real interest rate; however higher values of the Calvo parameter have a delaying effect on the behaviour of the real interest. In the baseline specification, the real interest rate falls by more on impact in the FA case and then goes back to its steady state value at a faster pace following an expansionary monetary shock. For higher values of the Calvo parameter, this process is delayed so that the real interest rate remains below its steady state level for a longer period.

Figure 4-3: Financial accelerator effect and price stickiness.



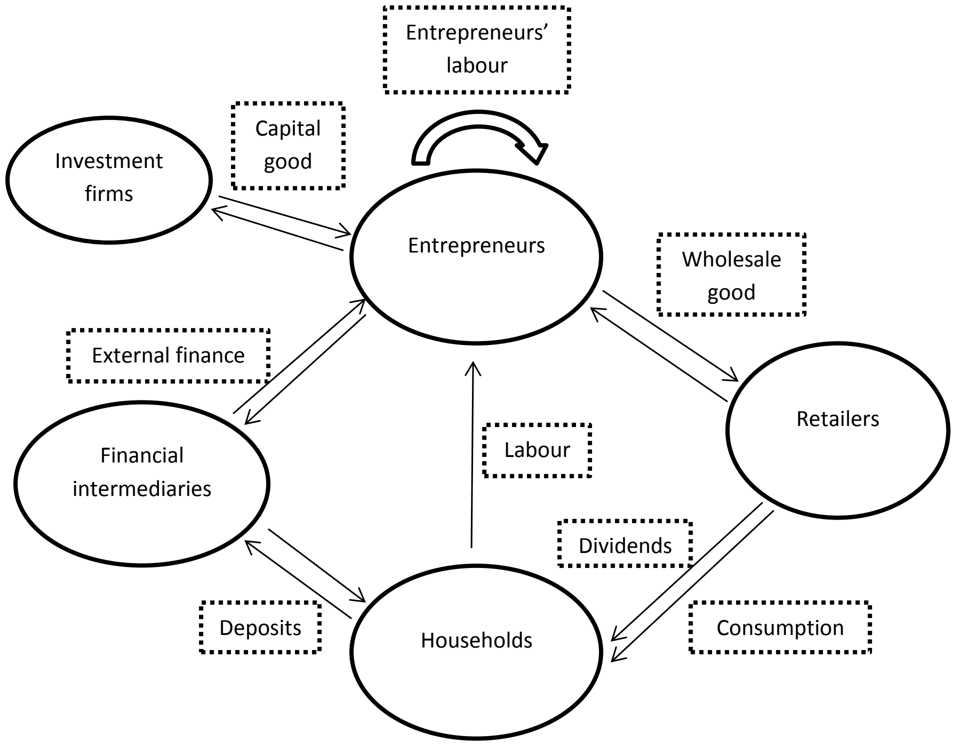


Figure 4-4: The environment

Chapter 5

Bank lending standards, monetary policy and the term spread

Abstract *Using US bank lending survey data, we revisit the previous evidence on the effect of monetary policy on bank risk-taking as measured by bank managers'* own ex-ante *risk perceptions (Maddaloni and Peydro, 2011; Buch et. al., 2014) by examining more thoroughly the role of the term spread in the context of a structural VAR analysis. We find robust evidence that an expansionary monetary policy in the US over the 15 years that preceded the 2007-2009 financial crisis led to an immediate tightening of bank lending standards followed by a subsequent loosening as predicted by recent theories of bank risk-taking. This relationship is also economically significant, the Fed Funds rate explaining a significant fraction of the forecast error variance of bank lending standards. In addition, an increase in the term spread leads to an immediate loosening of lending standards by banks, further accrediting the hypothesis that the effect of monetary policy on the quality of credit is in part supply-driven.*

Keywords: Bank risk-taking, Monetary policy, Financial frictions, SVAR

JEL Classification Numbers: E44, E51, E52, G21

5.1 Introduction

The 2007-2009 financial crisis, which saw a number of financial intermediaries go bankrupt or bailed out by governments in the US and in Europe, has generated a renewed interest in the determinants of bank risk-taking and especially the role of monetary policy. This new research agenda is tightly linked to the policy debates over whether conventional monetary policy should be concerned about the build-up of financial instability (Borio and Lowe, 2002) and over the need to link microprudential and macroprudential policies (Adrian and Shin, 2011). It also has implications for the regulation of international flows (Miranda Agrippino and Rey, 2014; Bruno and Shin, 2015).

Borio and Zhu (2012) defined the risk-taking channel of the monetary transmission mechanism as "the impact of changes in policy rates on either risk perceptions or risk tolerance, and hence on the degree of risk in portfolios, on the pricing of assets, and on the price and non-price terms of the extension of funding". In practice this notion has been applied to several distinct mechanisms. In a process reminiscent of the financial accelerator mechanism, an expansionary monetary policy may affect the valuation of assets, cash flows and profits, reducing moral hazard problems within the financial sector (Jimenez et. al., 2014; Adrian and Shin, 2014) or relaxing constraints imposed by capital regulation (Adrian and Shin, 2010). This in turn is associated with an increase in banks' leverage and a fall in risk premia or the pricing of risk (Adrian and Shin, 2011). A fall in the real risk-free rate may also spur a "search for yield" effect (Rajan, 2005), whereby investors with sticky target rates of return (such as institutional investors) shift away from safe assets and into riskier investments. On the theoretical side, Dell'Ariccia et. al. (2014) developped a formal banking model to account for this mechanism and show that a fall in the real interest rate could either give rise to a risk-shifting effect or a risk-taking effect on bank portfolio behaviour, depending on the state of the economy and possibly the timing of the effects, the relative importance of these raising therefore an empirical question.

On the empirical side, a major difficulty is the identification of credit supply and demand shocks, which, compounded by the need to take into account firms' and banks' heterogeneity, has made these effects difficult to test and quantify. The empirical literature on firms' and banks' balance sheet channels and on the bank lending channel of the transmission of monetary policy illustrates these difficulties[1]. The identification of a bank risk-taking channel poses additional challenges.

Most of the existing evidence relates changes in the policy rate to credit risk composition effects or search for yield effects over the decade preceding the Great recession. Jimenez et. al. (2014) and Ioannidou et. al. (2015) showed how micro-data at the bank-loan level can be used in identifying a risk-taking channel of monetary policy. Using two complementary approaches in the conditions of quasi-natural experiments in Spain and Bolivia, they find that following a fall in the policy rate, fragile banks are more likely to grant loans to riskier borrowers and this

[1]See Bernanke and Gertler (1995), Kashyap, Stein and Wilcox (1993, 1996), Oliner and Rudebusch (1996), Den Haan et. al. (2007).

extra risk-taking is included in the loan rate charged by banks on these loans. However, micro-based studies pose another string of identification problems such as the endogeneity of monetary policy and do not allow a clear cut quantification of these effects at the aggregate level. Time series methods are more appropriate to study these effects and allow an examination of the dynamic responses, taking into account the endogenous response of macroeconomic variables to monetary policy shocks on credit supply and risk-taking. The aggregate effects of changes in the policy rate on the composition of US bank loan portfolios are investigated in Buch et. al. (2014) who use a FAVAR approach and bank lending survey data. They find evidence that following an expansionary monetary shock small banks increase the relative size of riskier business loans on their balance sheets whereas loan rates on riskier loans fall.

Another strand of the literature has been studying the link between monetary policy rates and bank lending standards, which allows extending the sample period to start in 1990. Bank lending standards have been shown to be good predictors of real activity (Lown and Morgan, 2006)[2]. Maddaloni and Peydro (2009) show that lower policy rates in Europe and the US lead to a loosening of lending standards using panel regression analysis. Importantly, they also find that a steepening of the term spread (the spread between the 10-year Treasury note rate and the 3-month T-bill rate) in the US leads to a similar response. Adrian and Shin (2011) also postulate a link between low monetary policy rates, increases in the term spread and therefore in the profitability of bank portfolios and ultimately bank leverage.

The purpose of this chapter is to contribute to the empirical investigation of the risk-taking channel of monetary policy by studying directly the relationships between monetary policy, the term spread and bank lending standards in a VAR framework over the period 1990Q2-2007Q2. In particular, using impulse response analysis we investigate the time path of the response of the term spread and bank lending standards following a shock to the monetary policy rate. We find that a contractionary monetary policy leads to an immediate loosening and then a significant tightening of lending standards after two years. This finding is robust to different alternative identification schemes, different macroeconomic controls, using variables in first difference or in level and using an alternative sample period. In addition, we find evidence in favour of a chain of causation running from a fall in the policy rate to a steepening of the term spread and a loosening of bank lending standards. In addition, a tightening of lending standards leads to a fall in the growth rate of GDP, a relationship that is economically significant with bank lending standards explaining 10% of the forecast error variance of GDP growth at long horizons. Hence our findings are suggestive of an effect of monetary policy on real activity via its effect on bank lending standards.

The remaining of the chapter is organized as follows. In the next section, we review the empirical and theoretical literature on the risk-taking channel of monetary policy. In Section 5.3, we discuss the main identification problems and present the data. In section 5.4, we present the methodology, the model specification and the results. Section 6.5 concludes.

[2]Lown and Morgan (2006) estimate standard monetary VARs and show, using impulse response analysis and forecast error variance decomposition, that a tightening of bank lending standards leads to a decline in commercial loans and output, a relationship that is economically significant. Loan rates do not hold any additional explanatory power for these two variables.

5.2 Literature review

Borio and Zhu (2012) discussed three possible and complementary ways in which the risk-taking channel of the monetary policy transmission mechanism can manifest itself. According to the first view, changes in interest rates, as in financial accelerator models, affect asset valuations and profits of financial firms, leading to an increase in risk tolerance as measured by indices of stock market volatility such as the VIX or the risk composition of banks' loan portfolios. A second mechanism frequently discussed is a "search for yield" effect (Rajan, 2005) according to which investors with sticky targets of returns (such as institutional investors) would react to a fall in the yield of safe assets caused by an expansionary monetary policy by taking on riskier positions.

This literature is mostly empirical and more often than not only supported by anecdotical accounts or discussed as a possible extension to theories dealing mainly with bank portfolio choices (Adrian and Shin, 2011; Acharya and Naqvi, 2012). When this issue is tackled directly in the theoretical literature, the models point to more nuanced conclusions regarding the overall effect of changes in the risk-free interst rate on bank risk-taking (Dell'Ariccia et. al., 2014).

A third channel through which monetary policy may affect bank risk-taking works via central banks' communication policies and how they frame expectations about future policies (Fahri and Tirole, 2012). Evidence on these mechanisms is likely to be difficult to obtain as it would be difficult to identify in an empirical framework.

In the next section we will discuss the main identification problems raised in the empirical literature as well as their main results. In Section 5.2.2, we will review some of the theories that have been developpped to support these findings.

5.2.1 Risk-taking channel of monetary policy: identification problems and empirical evidence

Most of the empirical evidence in favour of a risk-taking channel of monetary policy linking low policy rates to increases in bank-risk-taking is concerned with testing one dimension of the first view mentionned above. In this section, we start by discussing two papers defending stronger versions of the risk-taking channel while clarifying the main identification problems using micro data (Jimenez et. al., 2014; Ioannidou et. al., 2015). We then discuss the closely related literature studying the link between monetary policy and bank lending standards (Dell'Ariccia et. al., 2013, 2013; Maddaloni and Peydro, 2011, Buch et. al., 2014; Loan and Morgan, 2006). Finally, we review the evidence consistent with the risk-taking channel more loosely defined using macro data.

Jimenez et. al. (2014), exploiting individual (firm-bank) loan level data from the credit register in Spain over the period 1988-2008, identify a risk-taking channel of monetary policy if a fall in the policy rate leads fragile banks (more likely to be ridden by moral hazard problems) to increase their relative supply of loans to riskier borrowers. This identification is achieved empirically by regressing banks' decision to grant a loan on a triple interaction term between changes in the short-term interest rate, bank capital-to-asset ratio and a firm credit risk measure[3], controlling for demand and balance sheet channels using time*bank and time*firm fixed effects. This empirical strategy relies crucially on the exploitation of information at the bank-firm level. Key to this identification scheme is the assumption that the bank-capital ratio is an adequate proxy for the strength of the agency problem within banks[4]. Mitigating the problem of the endogeneity of monetary policy, the policy rate set by the ECB is considered to be independent of developments in the Spanish economy.

They find that following a fall in the short-term interest rate, banks with lower capital ratios extend more credit to riskier firms both at the intensive and at the extensive margin.[5] This result is robust to introducing a triple interaction term with changes in long-term rates. They therefore provide evidence of a significant effect of short term rates on the (risk) composition of credit supplied by banks. Such an effect may not be identifiable using aggregate data: not finding a significant effect of monetary policy on an aggregate measure of bank risk-taking without taking into account banks' heterogeneity could still be consistent with this finding. However, uncovering a significant effect with aggregate data, and using appropriate controls, would represent a stronger result.

Ioannidou, Ongena and Peydro (2015) identify a similar channel, using detailed credit register data[6] from Bolivia between 1999 and 2003, a period during which the Bolivian peso was pegged to the US dollar and the banking system was largely dollarized[7]. This paper complements the approach taken by Jimenez et. al. (2014) by focusing on loan risk premia as an alternative method to distinguish loan supply from demand effects and risk-taking. To identify the impact of a lower Feds Funds rate on banks' risk taking, information about the contemporaneous changes in loan risk premia is needed: intuitively if the issuance of riskier loans is demand-driven, then it is more likely to be associated with a rise in loan premia (the price of risk). The result may also be driven by an increase in the demand for loans by riskier borrowers. To control for this possibility, the authors use a number of controls on firms', bank-firm relationships, banking market conditions and other macroeconomic factors. A further problem is that banks may adjust for the additional risk via changes in other terms of lending such as maturity and collateral value. Using such information, the authors estimate the expected return of loans at origination and examine whether it is negatively related to changes in the Feds Funds rate (which is more likely to be the case if the increased loan riskiness is demand-driven).

Their results are supportive of a chain of causation that runs from a fall in the short-term interest rate to increases in both ex-ante and ex-post bank risk-taking[8]. In addition, this effect

[3]Measures of credit risk considered are the number of doubtful loans firms have contracted over the last 4, 3, 2 and 1 year and in the next 1 and 2 years.

[4]Holmstrom and Tirole (1997) provide one such motivation.

[5]In a first set of regressions, the dependent variable is the growth rate of the balance of credit a particular firm has at a particular bank; in a second set of regressions, the dependent variable is binary, indicating whether loan credit was ended conditional on having a positive credit balance in the previous period; in a third set of regressions the dependent variables are binary, indicating whether the loan was granted in the case of new loan applications and conditional on the new loan being granted, the logarithm of the amount committed.

[6]In particular, the data comprises loan amounts, interest rates, maturities and collateral values of loans.

[7]This setting means that the results are less likely to be biased by the endogeneity of monetary policy.

[8]They first estimate probit models with an ex-ante measure of loan riskiness and find that a lower Fed funds rate at the time of loan origination is associated with banks issuing more loans to less creditworthy borrowers.

appears to be stronger for banks more likely to suffer from moral hazard problems, ie those holding a greater share of bad loans on their balance sheets and with lower capital ratios (only for the ex-post measure of risk-taking). However, a fall in the Fed funds rate occurring after the loan is committed tends to reduce the loan hazard rate. Finally, this additional risk-taking by banks does not appear to be reflected in higher loan rates and expected returns. On the contrary, loan rates appear to be negatively related with the increase in the hazard rate due to a fall in the Fed funds rate at the time of origination. This finding supports the hypothesis of an increase in banks' risk appetite as opposed to an increase in the demand for risky loans.

There is also some evidence that increased bank risk-taking is demand-driven. For example, Dell'Ariccia et. al. (2012) study the US mortgage market in the run-up to the 2007-2009 financial crisis using loan-level applications and find, after controlling for a number of supply-side factors and loan applicant characteristics as well as potential biases, that subprime loan applications were less likely to be denied in areas that saw a larger volume of applications. They interpret this result as evidence that demand side factors were partly driving the lax lending standards observed during this period. This finding is in line with theoretical papers predicting that an increase in the pool of borrowers may reduce adverse selection problems in the banking sector and induce banks to relax lending standards (Dell'Ariccia and Marquez, 2006). This result does not, however, hold for prime loan applications. Given the relatively small size of the subprime mortgage market compared to aggregate loan volumes, the macroeconomic significance of this effect comes from its potential implications for the overall fragility of the financial system. Although other segments of the loan market may differ substantially from the mortgage loan market[9], this finding suggests that demand effects may play an important role in explaining the supply of credit extended by banks.

Next, we discuss the papers using lending survey data to investigate the determinants of bank risk-taking[10]. Contrary to the previous studies, bank lending surveys convey information about bank managers' perceptions of the riskiness of the loans they accepted to take on (Buch et. al., 2014) and provide information about new loans.

Dell'Ariccia, Laeven and Suarez (2013) conduct a study of the risk-taking channel in the US[11] using loan level data over the period 1997-2011. Bank risk-taking is proxied by an ex-ante measure of loan risk assigned by the bank to each loan (rated on a scale from 1 to 5). This measure for each loan granted in each period is regressed on the real risk-free interest rate, a measure of the capital ratio of the bank and a set of controls. A crucial assumption supporting the identification approach is that the policy rate was not set in reaction to adverse developments in the financial system during this period. However, if loans become, ceteris

They also estimate a proportional hazard model (time-varying duration model) where the time to default or repayment of each individual loan (ie an ex-post measure of loan riskiness) is regressed on the levels of the Fed Funds rate at the time the loan is committed by the bank and over the life time of the loans. They find that a lower Fed funds rate at the time of loan origination is associated with a higher future probability of default.

[9] In terms of maturity, collateralization and interest rates.

[10] For the US bank lending survey data is collected by the Federal Reserve Board and either comes from the Senior Loan Officer Opinion Survey on Bank Lending Practices (SLOOS) which provides information on lending standards from 1992 onwards or the Survey of Terms of Business Lending (STBL) starting in 1997.

[11] The banking sector in the US accounts for a relatively lower share of the total amount of funding available to finance real investment projects in the economy.

paribus, more risky during recessions, when monetary policy is expansionary, it would not be possible to identify, using this framework, whether the increase in the riskiness of the loans is due to an increase in banks' risk appetite or an increase in firms' demand for risky loans even if some macroeconomic controls are introduced into the regression. To control for this issue, the regressions are run on subsamples comprising states with income less correlated with nationwide conditions, using the fact that monetary policy is set with respect to the overall macroeconomic conditions in the US.

The results strongly support the existence of a significant negative relationship between loan risk ratings and the Fed Funds target rate. The paper further investigates the relationship between increased bank risk-taking following a fall in the policy rate and banks' capital structure by introducing an interaction term between the short term rate and the capital ratio of the bank. They find that banks with higher capital ratios (less levered) are more likely to grant riskier loans following a fall in the policy rate[12]. Finally, following longer periods without interest rate rises, banks make riskier loans.

Maddaloni and Peydro (2011) examine the relative importance of low short-term rates, low long term rates and too low policy rates for a long period of time[13] in explaining changes in bank lending standards in the Euro area and the US. They find that low (short-term) policy rates for extended periods of time lead to a softening of lending standards and that this effect is higher, the higher the issuance of securitized assets and the weaker bank capital supervision (for mortgage loans in the Euro area). In contrast, lowerer long-term rates and a flattening of the term spread (in the US) lead to a tightening of lending standards.

Buch et. al. (2014) are able to exploit 140 aggregate time series from the STBL accounting for loan volumes, loan rates and other terms of lending (such as collateralization and maturity) for 4 different risk categories of C&I (commercial and industrial) loans and 3 different types of banks (large, small and foreign) by employing a FAVAR methodology. Changes in bank risk-taking are identified as shifts in the risk composition of the credit supplied by banks that are not compensated by movements in loan risk premia, holding constant the quality of borrowers. Given that riskier loans command a higher risk premium, the specific mechanism tested here is that of a search for yield.

They find significant evidence of such a channel operating for small banks (and foreign banks during the period where the policy rates were "too low for too long"), which may be due to the fact that they face more severe agency problems[14]. In particular, impulse responses to an expansionary monetary policy shock leads small banks to increase the supply of credit to high risk borrowers relatively to minimal risk borrowers and to lower loan spreads. In addition loan spreads to high risk borrowers fall significantly more than for minimal risk borrowers. These results support the hypothesis that the increased bank risk-taking is supply rather than demand-driven (or due to increased loan demand by riskier borrowers)[15]. Finally, these effects

[12]This result appears to be opposite to the findings of Jimenez et. al. (2013) and Ioannidou et. al. (2008).

[13]Negative Taylor rule residuals (the difference between the effective Fed Funds rate and the Taylor rule rate) directly translate the idea of a "too low policy rate".

[14]They set higher loan rates and larger spreads between rates on high risk and low risk loans.

[15]However, the fact that the response of loans to minimal risk borrowers is insignificant and for small banks

are associated with small banks reducing the maturity and increasing collateral requirements for loans to riskier borrowers, thus mitigating in part the cost of higher credit risk.

Lown and Morgan (2006) differ from the previous studies in terms of the sample coverage (which roughly covers the period 1967-2000) and the questions they address. Their focus is on the role of lending standards as a non-price term of lending in explaining changes in credit and in real activity. In particular they interpret the significant predictive power of changes in lending standards for these variables as indicative of a potent role for financial frictions that materializes through changes in the availability of credit (or rationing) rather than through prices (loan rates do not appear to explain loan volumes as well). In other words, their question is whether credit supply shocks can be better identified by changes in lending standards rather than by changes in loan rates. This paper points to the absence of a downward-sloping demand for loans relating higher loan rates to lower loan volumes. Although this suggests an important role for the non-price terms of lending set by banks, it renders the identification of bank risk-taking more difficult. Finally, Lown and Morgan (2006) do not find any significant explanatory power of changes in the Feds Funds rate for lending standards.

Finally, some studies at the macro level find results consistent with the risk-taking channel loosely defined. Den Haan et. al. (2007) examine the effect of monetary policy on bank loan portfolios and find that C&I loans increase following a monetary contraction and decrease during a non-monetary contraction, which may be explained by the relatively less risky nature of such loans (see Section 5.3.1). Bekaert et. al. (2013) find robust evidence linking expansionary monetary policies to falls in the option implied volatility of equity (VIX index). Adrian and Shin (2011) and Adrian, Estrella and Shin (2010) use a structural VAR framework including shadow banks' asset growth and the change in the VIX index as a measure of 'risk' to identify a risk-taking channel of monetary policy. Their identification strategy relies on tracing the effect of changes in the short-term rate on the term spread, banks' net interest margin (an indicator of the profitability of banks' maturity transformation activity) and the growth of their balance sheets. These two studies explicitly link changes in the policy rate to changes in the term spread and ultimately in the perceptions of risk. Bruno and Shin (2015) and Miranda Aggripino and Rey (2014) link changes in the short-term rate to changes in the leverage of the banking sector and cross-border flows, thus uncovering a global financial cycle. Angeloni et. al. (2015) link changes in the stance of monetary policy to more general measures of bank risk such as asset risk, funding risk and the volatility of bank equity.

5.2.2 Risk-taking channel of monetary policy: theory

In this section, we discuss two theoretical models that have been influential in thinking about the nexus between monetary policy rates (or rshort-term risk-free rates), banks' leverage and bank risk-taking. Dell' Ariccia et. al. (2014) develop a banking model, which can be seen as the most parsimonious framework to think about bank risk-taking behaviour. Adrian and Shin

negative while loan spreads fall suggest that low risk borrowers may be reducing their demand for bank loans. Note that, contrary to high risk borrowers, minimal risk borrowers have access to alternative sources of credit at favourable terms by definition.

(2014) borrow from both the banking literature and the asset pricing literature to develop a theory of bank leverage and of the pricing of risk.

Dell'Ariccia et. al. (2014) construct a simple model in a static setting in which banks' choices are the outcome of a profit maximization problem in an environment characterized by limited liability, a downward-sloping loan demand schedule and a perfectly elastic or exogenous supply of deposits. Banks face a trade-off between a higher degree of monitoring of their loan portfolio which leads to a lower probability of default on the one hand and lower monitoring costs on the other hand[16]. In addition to the loan rate, they choose the riskiness of their portfolio of assets (which is inversely proportional to the degree of monitoring) for which they need to compensate depositors in the form of a risk premium over and above the risk-free rate.

Following a fall in the real risk-free rate which lowers banks' cost of funds and increases intermediation spreads, the expected marginal benefit of lending in case of success increases and induces profit-maximizing banks to increase monitoring (and marginal cost) thus reducing risk-taking (a risk-shifting effect). To the extent that a fall in the risk-free rate also reduces the optimal lending rate it may also give rise to a portfolio reallocation effect working in the opposite direction. In this case, the fall in the loan rate reduces the bank's profit, which, given limited liability, makes it more profitable to lower monitoring costs even at the expense of a higher default probability and premium on deposits, thereby increasing risk-taking (a risk-taking effect). The overall effect will depend on the balance of these two forces, although the first one is more likely to hold in the short-run and when leverage is high (or for less capitalized banks)[17].

However allowing banks to adjust their leverage leads to an unambiguous increase in bank risk-taking and leverage. This is because in this case the risk-shifting effect is neutralised by a leverage effect. On one side the risk-shifting effect reduces the bank agency problem and as a result depositors demand a lower risk premium. As deposits become relatively cheaper, banks substitute deposits for capital for each unit of asset on their balance sheet. On the other side this rise in leverage induces banks to take more risk because of limited liability and commands a higher risk premium. Overall, following a fall in the risk-free rate, banks trade off a greater expected return and upside risk for a higher risk premium on the funding side.

Adrian and Shin (2014) explored the risk associated with the funding side of banks' balance sheets, in relation to their leverage decision, taking stock of the impact of new bank capital regulations on banks' behaviour. In their model banks finance their asset purchases through collateralized short-term debt and equity. Limited liability implies bank owners are granted a put option by their creditors increasing in value in the volatility of their assets. Thus, as in Holmstrom and Tirole (1997), they have an incentive to invest in inefficient and riskier assets. The optimal contract between bank shareholders and depositors implies an upper bound on the level of bank leverage and the bank probability of default. Moreover, as measured risks evolve over the financial cycle, they show that intermediaries adjust their leverage as if their aim was to maintain a constant the probability of default, thus formalizing the idea of banks

[16]Alternatively, the trade-off may be recast in terms of a risk-return trade-off.

[17]Because those banks are more sensitive to changes in the cost of funds.

managing their balance sheet size subject to a Value-at-Risk rule[18]. In particular, leverage is procyclical, that is, it is lower when risks are high (bust times) and agency problems induced by limited liability are stronger (implied put options are more valuable). To the extent that monetary policy has an effect on the strength of the VaR constraint, it will have an impact on bank leverage.

Adrian and Shin (2011) and Shin (2010) developed a simple asset portfolio model with leveraged and passive investors to explore further the asset pricing implications of these mechanisms. The former are risk-neutral but face a Value-at-Risk constraint which limits their portfolio choice and therefore their leverage so that they cannot default with a constant probability. The latter behave like mean-variance investors. For a fixed stock of risky securities, a relaxation of the VaR constraint due to a fall in risk perceptions or a rise in the value of their assets allows leveraged institutions to expand their balance sheets, increasing their share of credit relative to passive investors and leading to a fall in the market-wide risk-premium[19]. In the medium run, due to their increased risk-bearing capacity, leveraged investors may seek to expand the pool of borrowers receiving credit in the economy. Thus, the model provides a rationale for a risk-taking channel of monetary policy working through the balance sheets of leveraged financial intermediaries to influence both the market-wide risk premium and the supply of credit.

5.3 Empirical strategy

5.3.1 Main identification problems

The main objective of this chapter is to examine how changes in the monetary policy rate affect bank lending standards and whether changes in the term spread plays a role in these processes. Thus we propose to complement the intuitions contained in Loan and Morgan (2006), Maddaloni and Peydro (2009) and Adrian and Shin (2011) by investigating directly the dynamic interactions between lending standards, monetary policy and the term spread. Existing theories suggest (at least) two channels through which the term spread may influence bank risk-taking. First, by increasing the profitability of the maturity transformation activities of banks, a widening of the term spread may lead to increased leverage (Adrian and Shin, 2011). Second, a widening of the term spread may give rise to both a risk-shifting effect (by making it more valuable for banks to ensure that borrowers do not default) and a leverage effect (by inciting banks to increased debt borrowing through a dampening of agency problems) leading them to take more risks (Dell'Ariccia et. al., 2014). We use the *Senior Loan Officer Opinion Survey on Bank Lending Practices* (SLOOS) dataset collected by the Federal Reserve Board. Table 5.1 describes the variables of interest. In this section, we discuss the main identification problems.

[18]Such a rule can also be motivated by the role of bank capital regulation.

[19]These asset pricing consequences have been explored empirically by Adrian, Etula and Muir (2014). They show that banks' leverage, which is inversely related to their marginal value of wealth is an important determinant of the stochastic discount factor and prices a wide range of financial assets in the US.

One potential problem identified by the literature discussed in the previous section is the endogeneity of monetary policy with respect to real activity. On one hand, loan risk may be expected to be higher in a downturn. On the other hand, monetary policy is likely to be expansionary during a recession. This may lead to a spurious negative relationship between the two variables and provides a strong argument for using a VAR framework including all three variables.

Second, to the extent that a loosening of lending standards allows riskier firms to obtain credit, a significant negative relationship between the monetary policy rate and bank loans to riskier firms or tightened lending standards may be explained by the presence of a broad credit channel (Oliner and Rudebusch, 1996) or a flight to quality (Bernanke et. al., 1996). According to this view, a contractionary monetary policy shock leads to a deterioration of agency problems affecting more strongly the most vulnerable (high risk) firms and as a result reduces the share of credit received by these firms in the economy. This mechanism is consistent with a shift in the aggregate mix between market and bank credit (Kashyap et. al., 1993) because firms suffering from more severe agency problems are more likely to be bank-dependent, but there is also some evidence that it is at work within the banking sector[20]. In addition, the broad credit channel predicts that the external finance premium should increase for all firms following a contractionary monetary policy although it may be more pronounced for firms more exposed to agency problems. Finally, the early literature on the role of financial frictions for business cycle fluctuations (Stiglitz and Weiss, 1981; Bernanke and Gertler, 1989) predicts an improvement in the overall credit quality of borrowers in the economy following a expansionary monetary policy shock. Important implications are that for small/riskier firms, lending standards may be expected to be tightened more strongly and the external finance premium to rise by more. In the likely case where current macroeconomic conditions have no effect on the risk rating of a firm (reflecting for example the inherent heterogeneity of firms in terms of size, collateral value, exposure to the macroeconomic environment and access to alternative source of funding), a change in the risk composition of bank credit may therefore reflect a firm balance sheet channel. This makes it very difficult to disentangle a firm balance sheet effect from a bank risk-taking effect empirically.

Third, a negative relationship between loan risk and the monetary policy rate does not necessarily reflect a change in banks' risk appetite as it could also be attributed to changes in the demand for riskier loans. To distinguish between loan supply and demand shocks, we need information about the simultaneous behaviour of loan rates. If the loosening of bank lending standards following an expansionary monetary policy is due to an increased demand for loans by less creditworthy borrowers (see Dell'Ariccia et. al., 2012), then we should observe a contemporaneous relative increase in the spreads between high and low risk loan rates. Conversely, to the extent that monetary policy does not alter the relative riskiness of the borrowers (or the distribution of risks across borrowers[21]), significant evidence in favour of the opposite effect

[20] See Bernanke, Gertler and Gilchrist (1996) for a review of the early literature.

[21] Similarly, Dell'Ariccia et. al. (2013) raise the possibility that the state of the business cycle might bias loan officers into over(under)-estimating loan riskiness during a contraction(expansion).

may be interpreted as a shift in the pricing of risk by banks (a credit supply shock).

Fluctuations in the volume of C&I (commercial and industrial) loans[22] have been investigated in the context of the transmission of monetary policy (Bernanke and Gertler, 1995; Den Haan et. al., 2007) where their "non-standard" response has been noted. In particular, using a comprehensive dataset from the Call reports, Den Haan et. al. (2007) find that banks tend to increase the volume of C&I loans on their balance sheet following a contractionary policy but to decrease it following a non-monetary downturn. They explain this finding via the hedging of maturity risk and the effect of short-term interest rates on banks' profitability. Using data from the STBL based on a sample of around 400 banks but uniquely providing information on *new* loans and including credit lines, Buch et. al. (2014) do however find that an expansionary monetary policy leads to an increase in C&I loans. C&I loans also feature in the literature on the role of banks as liquidity providers for borrowers due to the important fraction of those loans that are issued under commitment by banks (Gatev and Strahan, 2006; Ivashina and Scharfstein, 2010). According to this view, an additional dollar of credit supplied by banks may also represent the decision of borrowers to draw down their credit line in response to the deterioration of funding conditions. One advantage of the SLOOS data is that it uniquely contains information on new loans and therefore provides a way of disentangling loan supply changes due to credit line drawdown from those due to the granting of new loans.

5.3.2 Senior loan officer opinion survey data

The SLOOS is collected on a voluntary basis across a sample of up to 60 US domestic banks and up to 24 large US branches of foreign institutions. Domestic banks are selected on the basis of their size, location and are mutually independent of each other. In 2011, their assets represented 69% of the total assets of domestic banks. The same year, sampled foreign banks held 55% of all C&I loans held by foreign institutions in the US[23]. Questions in the survey are mostly qualitative and the aggregate results are presented as the percentage of respondents reporting having taken a given action regarding the terms of C&I lending offered over the previous quarter.

Our main measure of interest is the net percentage of banks declaring tightening standards constructed as the ratio of the difference between the number of loan officers reporting tightening and those reporting easing over the total number of respondents[24]. This question started to be asked on a regular basis in 1990 so our sample starts in the second quarter of this year. We take this measure as the empirical counterpart to the degree of bank loan monitoring in the model of Dell' Ariccia et. al. (2014), subject to the limitations discussed in section 5.2.1. In other words, as banks tighten lending standards, they increase their screening activity on the

[22] C&I loans tend to be short-term loans, are more likey to be issued under commitment (ie to involve a credit line) and collateralized.

[23] Information regarding the survey can be found at http://www.federalreserve.gov/boarddocs/SnLoanSurvey/about.htm.

[24] More specifically, banks answer the following question: "Over the past three months, how have your bank's credit standards for approving applications for C&I loans or credit lines - other than those to be used to finance mergers and acquisitions - to large and middle market firms changed?" (see Lown and Morgan (2006) and Maddaloni and Peydro (2011) for more detail).

new loans they issue, which reduces the default probability on these loans, hence a reduction in bank risk-taking. Banks also report whether they increased the spread between loan rates and their cost of funds and whether they observed a stronger loan demand from firms.

Figure 5-1 plots the three time series for both large and small firms. The series for the two categories of firms follow each other closely, the most notable difference being that changes in the spread tended to be less volatile for small firms until the early 2000s. Although the 1990-1991 and 2001 recessions appear to have been preceded by large spikes in lending standard tightenings, we do not observe this pattern in the run-up to the 2007-2009 recession. Rising intermediation spreads only preceded the 2001 recession. Finally loan demand falls just before recessions, then rises at the onset of recessions and reach its trough towards the end. We select changes in lending standards to large and middle-market firms as the (inverse) measure of bank risk-taking.

5.4 Empirical analysis

5.4.1 Methodology

We consider the following VAR model:

$$Y_t = BY_{t-1} + CX_t + \epsilon_t \tag{5.1}$$

where $Y_t = [\Delta GDP_t, \Delta GDPDEF_t, FF_t, TERM_t, LS_t]$ is the vector of endogenous variables and X_t is a vector of exogenous variable including a constant term. ΔGDP_t is the log difference of real GDP, $\Delta GDPDEF$ is the log difference of the GDP deflator, FF is the effective Fed Funds rate, $TERM$ is the difference between the 10-year Treasury bond rate and the 3-month Tbill rate and LS is the net percentage of banks tightening their lending standards for loans to large firms. Stability tests are reported in Table 5.3 and indicate that the null of no breakpoint in all equations is rejected in all equations. The likely breakpoint varies across equations, although the most frequent date (and the one given for the full VAR) appears to be the first quarter of 2008; in view of these results, we end our sample in the fourth quarter of 2007[25] so that we restrict the sample to the pre-crisis and pre-zero lower bound period. The AIC and SIC criteria select lag lengths of 2 and 1 respectively[26]. All eigenvalues are inside the unit circle, implying that the system is stable.

5.4.2 Preliminary results

Var coefficients and t-stats for this specification estimated with one lag are presented in the upper left corner of Table 5.4 (Panel A) alongside those of two alternative specifications, one with the term spread replaced by the 10-year Treasury bond rate (model 2) and the other

[25] Bekaert et. al. (2013) and Bruno and Shin (2015) use similar end dates to conduct their analysis.

[26] LM tests for residual autocorrelations reveal that with 1 lag we cannot reject the null of serial correlation in the residuals for the Fed Funds and the Term spread equations at the 5% significance level. Including 2 lags removes all serial correlation.

without the term spread (model 3). In addition to providing evidence on the magnitude and the direction of the effects, because only one lag of the variables is included in the specifications, the t-stats can be given the interpretation of Granger causality tests. A rise in the Fed Funds rate is associated with a tightening of lending standards. However, a rise in the 10-year Treasury bond yield leads to a loosening of lending standards. This suggests that the effect of the short-term rate on lending standards does not operate via its effect on the long term rate (Maddaloni and Peydro, 2011). When including the term spread, the relationship between the Fed Funds rate and bank lending standards becomes insignificant whereas a steepening of the yield curve leads to a significant loosening of lending standards. In addition, the term spread equation (column 4 of model 1) indicates that the lagged Fed Funds rate is negatively related to the level of the current term spread, although the relationship is not significant. This is consistent with the evidence provided by Adrian and Shin (2011) who show that over the period considered, a rise in the Fed Funds rate led to a one-for-one fall in the term spead in the short-run. Taken together, these results suggest a role for the term spread in the transmission of monetary policy to bank lending standards.

When replacing the Fed Funds rate as the measure of the stance of monetary policy by Taylor rule residuals (Panel B), the latter appear to be marginally significant only when we include the long term rate. Taylor rule residuals are computed as the difference between the Fed Funds rate and the Taylor rule rate and thus negative Taylor rule residuals represent an excessively expansionary monetary policy. Using this specification, a rise in the term spread still leads to a significant loosening of lending standards.

Banks borrow short term and lend long term, therefore the term spread may indicate the profitability of the bank maturity transformation activity at the margin (Adrian and Shin, 2011). To investigate further how the term spread affects bank lending standards, we examine alternative specifications with in turn the net interest margin of commercial banks (NIM) and the growth rate of commercial banks' equity (ΔBE). If the transmission of monetary policy shocks to bank lending standards operates through changes in the profitability of banks, via the term spread, then we should find that other measures of the profitability of banks' long term investments affect lending standards in a similar fashion. Coefficients estimates for the lending standard equation only are presented in Table 5.5[27]. If the term spread is capturing the profitability of banks' lending activity, then a rise in the term spread would lead to a rise in the net interest margin and in turn a rise in the NIM would lead to a loosening of lending standards. We find some evidence in favour of this hypothesis in the model with the Fed Funds rate. The lagged NIM is negatively correlated with current lending standards (although this relationship is not significant) while the coefficient on the term spread becomes insignificant. Coefficients in the model with the change in equity are insignificant and have the wrong sign whereas the term spread remains significant.

We also investigate whether other measures of bank risk-taking may explain lending standards by examining specifications with the growth rates of "core liabilities" of commercial

[27] The first column gives the coefficient estimates for the model with the Fed Funds rate and the second column gives the coefficient estimates for the model with Taylor rule residuals.

banks (ΔCL) as a proxy for bank funding risk (Angeloni et. al., 2015) and the growth rate of ABS commercial paper issuance as a proxy for securitisation (ΔSEC) (Maddaloni and Peydro, 2009). A rise in the lagged growth rate of "core liabilities", which is a measure of the banking sector's short-term wholesale funding and is often associated with leverage (Angeloni et. al., 2015; Adrian and Shin, 2010), leads to a tightening of current bank lending standards. This runs counter the argument that an increase in short-term debt that is due to a rise in bank leverage corresponds to an increase in bank risk taking. An increase in the lagged growth rate of securitisation leads to a loosening of bank lending standards. This is consistent with theories predicting that securitisation allowed banks to increase their loan supply and reduce loan monitoring by passing some of the risk attached to it on to off-balance sheet vehicles. Finally we include the growth rate of real C&I loans (ΔCIL) and the net percentage of banks reporting increasing the spread between loan rates and their funding costs ($SPREAD$). An increase in the lagged growth rate of C&I loans leads to a loosening of current bank lending standards, suggesting that the margin of adjustment of loan levels is through a loosening of lending standards[28]. A rise in the spread variable leads to a tightening of lending standards. However there is not clear theoretical rationale about the sign of this coefficient[29]. None of the coefficient estimates discussed in this paragraph appear to be significant whereas the term spread remains significant in all specifications.

Finally, we add the net percentage of banks reporting a stronger demand for loans ($DEMAND$) in our baseline specification. Coefficient estimates and t-stats presented in Table 5.6 suggest that demand effects may play an important role in banks' risk-taking choices. An increase in the demand for loans leads to a loosening of lending standards. In addition there is evidence of reverse causality, a tightening of lending standards leading to a fall in the demand for loans. These effects are significant and suggest a role for feedback loops. Furthermore, there is a strong effect of GDP growth on the demand for loans, pointing to a second feedback effect, from GDP growth to lending standards via the demand for loans.

5.4.3 Impulse response functions and forecast error variance decomposition

Next we examine how these effects may play out over time by looking at the impulse response functions over a seven year horizon. We consider two methods of identification of structural innovations given the reduced-form residuals obtained by estimating equation (5.1) with two lags of the endogenous variables[30]. First, we employ a Cholesky decomposition of the variance-covariance matrix of reduced-form residuals and impose short-run restrictions by choosing the ordering $Y_t = [\Delta GDP_t, \Delta GDPDEF_t, FF_t, TERM_t, LS_t]$, implying that the policymakers react contemporaneously to GDP growth and inflation news but not to the term spread and lending standards as is usually assumed (Christiano et. al., 1999). However bank survey data is collected

[28] Lown and Morgan (2006), using a measure of loan supply in level provide evidence in favour of a significant relationship in the opposite direction.

[29] In the model of Dell'Arriccia et. al. (2014) the overall relationship between the intermediation spread and bank risk-taking is ambiguous.

[30] In order to identify the system of 5 variables completely, we need to impose 10 restrictions.

by the Federal Reserve Board in order to inform policy makers and therefore there is no clear rationale for not including lending standards in the information set of the monetary authority[31]. Second, we use a mix of short-run and long-run restrictions. Instead of having GDP growth not react contemporaneously to monetary policy, we impose the long-run restriction that monetary policy does not have a long-run effect on the rate of growth of GDP (see Appendix B for more details about these identification schemes).

The upper panels of Figure 5-3 present the impulse responses of the measures of real activity to a positive one standard deviation shock to the Fed Funds rate. The responses of GDP growth are either insignificant (short-run restrictions) or indicate a positive significant effect on impact (long-run restriction and short-run restrictions)[32]. The response of inflation is either insignificant or positive, therefore indicative of a price puzzle.

The lower panels of Figure 5-3 present the impulse responses of the term spread and lending standards. Following a contractionary monetary policy the term spread falls and this dampening persists for about two years. Lending standards are loosened on impact and then are tightened, reaching a peak after about two years and significantly persisting for about four years. The initial loosening of lending standards may be due to the prevalence of a risk-shifting effect, a rise in the cost of funds making it less attractive to screen borrowers and as a result bank risk-taking increases. As loan rates start increasing gradually this effect loses its strength and banks find it more profitable to monitor new borrowers more closely (Dell'Ariccia et. al., 2014).

Inspection of the impulse responses of the term spread suggests an alternative explanation. The impulse responses of lending standards to a shock to the term spread (Figures 5(g) and 5(h)) show that a fall in the term spread leads to an immediate tightening of lending standards. The flattening of the yield curve makes banks' portfolios of loans and securities, which are usually of a longer maturity than liabilities, less profitable and leads banks to reduce their leverage and risk-taking (Adrian and Shin, 2011; Dell'Ariccia et. al., 2014).

Figures 5(e) and 5(f) show that a reverse causation effect may also be at play, with a tightening of lending standards leading to a rise in the term spread in the medium term. One explanation is that policymakers care about conditions in the loan markets and react to a tightening of lending standards by lowering the policy rate. On impact this leads to an increase in the term spread which tends to ease somewhat lending conditions. This result does not change when we order lending standards before the Fed Funds rate, therefore making it part of the policymakers' information set (see Appendix A). The upper panels of Figure 5-4 examine the interactions between lending standards and GDP growth. A tightening of lending standard leads to a sharp but temporary contraction of output. An unanticipated expansion of output leads to a marginally significant tightening of lending standards after two years in the specification with only short-run restrictions. When restricting the long-run effect of monetary policy on output growth to be zero, it leads to an immediate tightening of lending standards,

[31] We also estimate impulse response functions using an alternative ordering placing lending standards in third position before the policy rate. Results are reported in appendix A.

[32] By ordering lending standards before the Fed Funds rate (appendix), we obtain a significant temporary fall of GDP growth as expected.

giving way to a significant loosening after two years and a half [33].

Finally, Figure 5-5 presents the forecast error variance decomposition of GDP growth and lending standards over the same horizon. The Fed Funds rate explains 20 to 30% of the variance of lending standards, implying an economically significant effect of monetary policy on bank risk-taking. In contrast, the term spread accounts for less than 5% of the variance of lending standards at any horizon. The fraction of the variance of lending standards accounted for by GDP growth and inflation is larger in the specification with the long-run restriction (7% and 10% respectively). In turn, lending standards explain about 15% of the forecast error variance of GDP growth.

5.4.4 Sensitivity checks

In this section, we conduct several robustness checks. First, we include in the analysis the net percentage of banks reporting stronger demand for C&I loans from the SLOOS dataset, to control for demand-side effects. Second, we investigate the interaction between monetary policy, lending standards and the VIX index. Third, we control for firm balance sheet effects by introducing a measure of the external finance premium. In all three specifications, a contractionary monetary policy leads to a tightening of bank lending standards, a result that remains economically significant with the Fed Funds rate explaining above 15% of the forecast error variance of lending standards at long horizons. Moreover, a tightening of lending standards leads to an immediate and significant fall in GDP growth in all three models, with lending standards explaining at least 10% of the error variance of GDP at long horizons.

Figures 5-6 to 5-8 show the impulse responses from a VAR specification where the term spread is replaced by the demand for loans. As expected, a monetary policy tightening leads to a significant fall in loan demand after three years. A tightening of lending standards leads to a fall in loan demand that lasts for about two years. A rise in loan demand leads to an immediate and temporary loosening of bank lending standards. Finally, loan demand explains about 10% of the forecast error variance of lending standards in the long run while 22% is still explained by the Fed Funds rate. This suggests a nonnegligible role for demand side factors in explaining changes in lending standards alongside supply side factors.

Figures 5-9 to 5-11 show the impulse responses from a VAR specification where the VIX index replaces the term spread and is ordered last. Consistent with Bekaert et. al. (2013) and Bruno and Shin (2015), an expansionary monetary policy leads to a significant fall in the VIX after three years. A tightening of lending standards leads to an immediate and significant rise in the VIX for about two years whereas the response of lending standards to a positive shock to the VIX is positive and significant only on impact. Miranda Agrippino and Rey (2014) and Bruno and Shin (2015) have shown how a rise in the VIX is associated with a fall in banks' leverage. Here we show that a rise in the VIX also leads to a tightening of lending standards,

[33] However, closer inspection of the impulse responses to an unanticipated rise in GDP (not shown) reveals a non-standard response of the Fed Funds rate which falls significantly, implying that policymakers would react to a fall in output by raising interest rates. We should therefore be cautious about interpreting the impulse responses with short and long run restrictions.

indicating that banks are less willing to take risks, as expected. In addition, in the long run the VIX explains about 7.5% of the error variance of lending standards. Our results also suggest the presence of a feedback loop from tighter lending standards to higher VIX index.

Figures 5-12 to 5-14 show the impulse responses from a VAR specification where the Baa-Aaa spread replaces the term spread and is ordered last. Contrary to the predictions of the firm balance sheet channel (Bernanke et. al., 1999), a contractionary monetary policy leads to a fall in the Baa-Aaa spread. However a rise in the spread leads to a tightening of lending standards after three years, as expected. As in the previous case we observe a large and significant effect of a tightening of lending standards on the spread. In addition, the Baa-Aaa spread explains a large fraction of the error variance of lending standards.

Finally, Appendix B shows additional robustness checks. The main results remain unchanged when we order lending standards before the policy rate, when we use a measure of output gap instead of the growth of real GDP, when we specify the variables in levels instead of first differences and finally when we conduct the analysis over the full sample.

5.5 Conclusion

We find robust evidence that an expansionary monetary policy in the US over the 15 years that preceded the 2007-2009 financial crisis led to an immediate tightening of bank lending standards followed by a subsequent loosening as predicted by recent theories of bank risk-taking. We cannot however conclude, within the empirical framework adopted in the chapter and due to data limitations, whether the loosening of lending standards is due to increased bank risk-taking or a higher demand for loans by riskier borrowers. This relationship is also economically significant, the Fed Funds rate explaining a significant fraction of the forecast error variance of bank lending standards. In addition, we find consistent evidence linking falls in the Fed Funds rate, to a widening of the term spread and the easing of credit conditions for businesses. This finding complements earlier results by Adrian and Shin (2011) and needs to be explained by the theories of the risk-taking channel of monetary policy. In view of the significant predictive power of bank lending standards for real activity, these results may also be linked to the well-known ability of the term spread to predict GDP (Adrian and Shin, 2011). Our results also suggest that demand side factors and firm balance sheet effects should not be neglected when evaluating the importance of changes in the policy rate on bank risk-taking.

Table 5.1: Bank risk-taking measures

Variable	Sample period	Obs
Senior Loan Officer Opinion Survey (SLOOS) on domestic bank lending practices		
Net % of banks increasing spreads of loan rates over banks' cost of funds to large and middle-market firms	1990Q2-2015Q2	101
Net % of banks increasing spreads of loan rates over banks' cost of funds to small firms	1990Q2-2015Q2	101
Net % of banks tightening standards for C&I loans to large and middle-market firms	1990Q2-2015Q2	101
Net % of banks tightening standards for C&I loans to small firms	1990Q2-2015Q2	101
Net % of banks reporting stronger demand for C&I loans from large and middle-market firms	1991Q4-2015Q2	95
Net % of banks reporting stronger demand for C&I loans from small firms	1991Q4-2015Q2	95

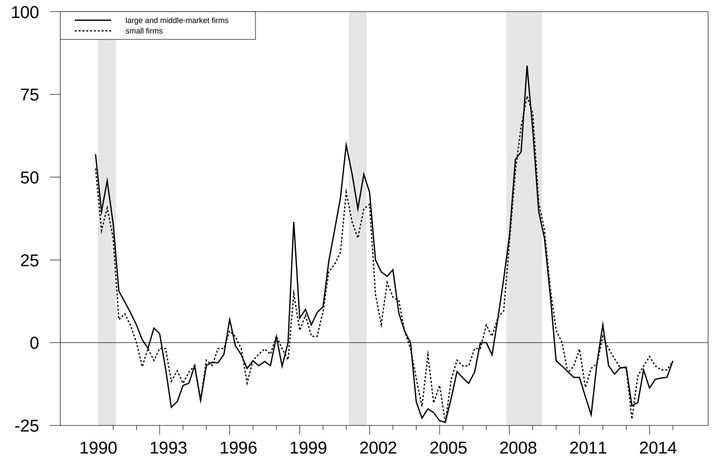
Note: The data is from the *Board of Governors of the Federal System* (Data download program).

Table 5.2: Macroeconomic data

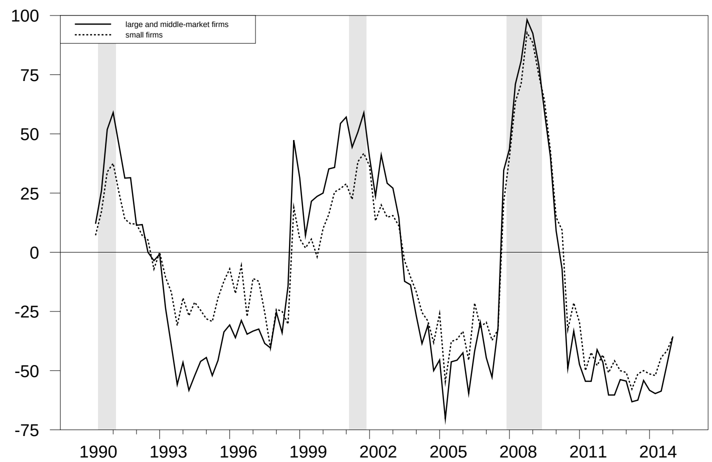
Variable	Empirical counterpart	label
Real activity	Real Gross domestic product	GDP
	Total payroll employment	EMP
	Industrial production index	INDPRO
	Chicago Fed National Activity index	CFNAI
	Output gap (HP detrended series of the log of real GDP)	YGAP
C&I loans	Total volume of commercial and industrial loans (all commercial banks)	CIL
	Bank prime loan rate	PRIME
	Delinquency rate on business loans	DEL
Inflation	Gross domestic product implicit price deflator	GDPDEF
Monetary policy stance	Feds Funds rate	FF
	Monetary aggregate M1	M1
	Taylor rule residuals (Difference between the Fed Funds rate and the Taylor rule rate)	TR
Interest rate variables	CP-Tbill spread (3-month)	CPSPREAD
	Term spread (10-year Treasury constant maturity rate minus 3-month T-Bill rate)	TERM
Measure of uncertainty	Chicago Board Options Exchange VIX index	VIX
Bank variables	Security Broker dealer leverage	LEV
	Core liabilities (sum of repo and large time deposits at commercial banks)	CL
	Bank equity	BE
	Net interest margin	NIM

Note: The data is from FRED. Variables are all quarterly (seasonally adjusted) data over the sample period 1990Q2-2015Q1. The Taylor rule rate is estimated using the following regression: $TR = \pi_t^a + NR_t + 0.5(\pi_t^a - \pi_t^T) + 0.5YG_t$ where π^a is the annual inflation rate, $NR = 2\%$ is the natural real interest rate, $\pi^T = 2\%$ is the target inflation rate and YG is the % deviation of real GDP from the real potential GDP (see Taylor (1993) and Bekaert et. al. (2013)).

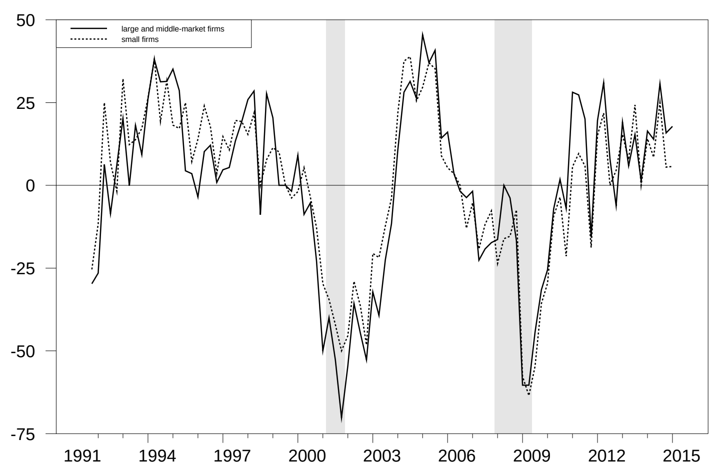
Figure 5-1: Plot of SLOOS series



(a) Net percentage of domestic banks tightening standards for C&I loans



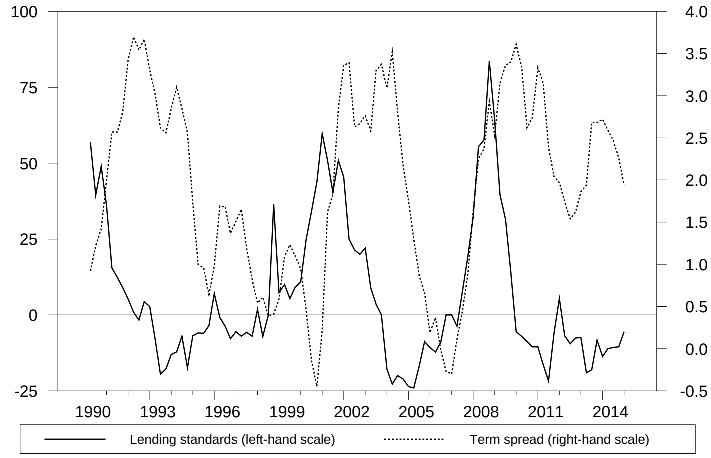
(b) Net percentage of domestic banks increasing spreads of loan rates over banks' cost of funds



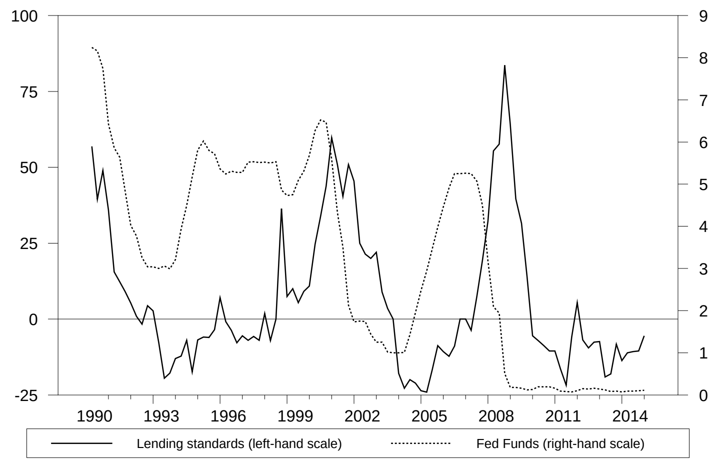
(c) Net percentage of domestic banks reporting stronger demand for C&I loans

Note: The data is from the *Board of Governors of the Federal System* (Data download program). Shaded areas represent NBER recessions.

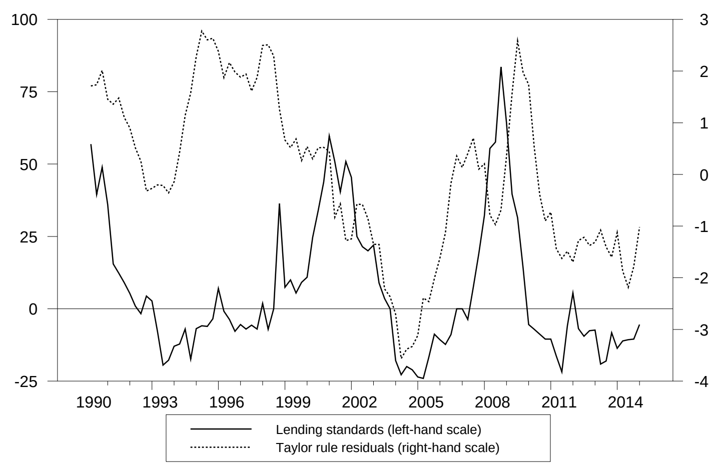
Figure 5-2: Monetary policy, the term spread and lending standards



(a) Lending standards and the term spread



(b) Lending standards and the Fed Funds rate



(c) Lending standards and Taylor rule residuals

Note: Lending standards are the net percentage of banks reporting tightening lending standards for C&I loans in the Senior Loan Officer Opinion Survey (Federal Reserve Board) from 1990Q2 to 2015Q1. The term spread is the difference between the 10-year Treasury constant maturity rate and the 3-month T-Bill rate. The Fed Funds rate is the quarterly average of the Effective Fed Funds rate. Taylor rule residuals are computed as the difference between the Effective Fed Funds rate and the Taylor rule rate (see Table 5.2 for detail of calculation).

Table 5.3: Stability tests

Sup-Wald stat (no break point in the estimated sample)			
Equation	Test statistics	p-value	breakpoint
ΔGDP	186.18	0.000	2006Q2
$\Delta GDPDEF$	97.80	0.000	2001Q3
S	297.65	0.000	2008Q2
FF	2465.75	0.000	2008Q1
$TERM$	2465.75	0.000	2008Q1
Full VAR	1958.66	0.009	2008Q1

Note: The VAR is estimated with 2 lags according to the AIC criterion. The Sup-Wald test allows for a general form of heteroskedasticity. P-values are asymptotic p-values (Hansen, 1996). The breakpoint is estimated as the date t at which the residual sum of squares (log determinant of covariance matrix) is minimized.

Table 5.4: Interest rates and lending standards

The table presents coefficient estimates of VARs estimated with 1 lag and a constant term (not reported) over the period 1990Q2-2007Q4 (70 obs). ΔGDP is the quarterly growth rate of real GDP. $\Delta GDPDEF$ is the quarterly change in the GDP deflator. FF is the Effective Fed Funds rate. TR are Taylor rule residuals computed as the difference between the Fed Funds rate and the Taylor rule rate (see Table 5.2 for details of calculations). $TERM$ is the difference between the 10-year Treasury bond rate and the 3-month-Tbill rate. LT is the 10-year Treasury bond rate. LS is the net percentage of banks reporting tightening lending standards for loans over the last quarter (SLOOS). *Panel A* reports coefficients with the Feds Funds rate as the measure of the stance of monetary policy. *Panel B* reports coefficients with Taylor rule residuals as the measure of the stance of monetary policy.

Panel A: Fed Funds Rate														
	Model 1					Model 2					Model 3			
	ΔGDP	$\Delta GDPDEF$	FF	$TERM$	LS	ΔGDP	$\Delta GDPDEF$	FF	LT	LS	ΔGDP	$\Delta GDPDEF$	FF	LS
ΔGDP_{t-1}	-0.018	0.076	0.314***	-0.178*	-1.122	-0.017	0.075	0.312***	0.136	-1.084	-0.018	0.074	0.311***	-1.321
	(0.892)	(0.703)	(0.001)	(0.097)	(0.654)	(0.895)	(0.705)	(0.001)	(0.205)	(0.664)	(0.891)	(0.705)	(0.001)	(0.607)
$\Delta GDPDEF_{t-1}$	-0.197**	0.522***	0.136**	-0.053	-1.583	-0.197**	0.522***	0.135**	0.063	-1.439	-0.197**	0.522***	0.134**	-1.695
	(0.012)	(0.000)	(0.014)	(0.391)	(0.281)	(0.012)	(0.000)	(0.015)	(0.315)	(0.327)	(0.011)	(0.000)	(0.015)	(0.261)
FF_{t-1}	-0.029	-0.035	0.963***	-0.0258	0.558	-0.027	-0.024	0.984***	0.001	2.868***	-0.029	-0.028	0.980***	1.718***
	(0.493)	(0.594)	(0.000)	(0.459)	(0.497)	(0.516)	(0.709)	(0.000)	(0.982)	(0.001)	(0.358)	(0.567)	(0.000)	(0.008)
$TERM_{t-1}$	0.000	-0.015	-0.037	0.933	-2.583**	-	-	-	-	-	-	-	-	-
	(0.999)	(0.873)	(0.405)	(0.000)	(0.035)	-	-	-	-	-	-	-	-	-
LT_{t-1}	-	-	-	-	-	-0.004	-0.008	-0.009	0.929***	-2.438**	-	-	-	-
	-	-	-	-	-	(0.942)	(0.925)	(0.822)	(0.000)	(0.031)	-	-	-	-
LS_{t-1}	-0.010***	0.000	-0.010***	0.009***	0.814***	-0.010***	0.000	-0.010***	0.000	0.804***	-0.010***	0.000	-0.010***	0.802***
	(0.002)	(0.947)	(0.000)	(0.000)	(0.000)	(0.002)	(0.958)	(0.000)	(0.884)	(0.000)	(0.001)	(0.959)	(0.000)	(0.000)
Panel B: Taylor rule residuals														
	Model 1b					Model 2b					Model 3b			
	ΔGDP	$\Delta GDPDEF$	TR	$TERM$	LS	ΔGDP	$\Delta GDPDEF$	TR	LT	LS	ΔGDP	$\Delta GDPDEF$	TR	LS
ΔGDP_{t-1}	-0.012	0.036	0.170	-0.176	-1.097	0.013	-0.014	0.135	0.142	-0.626	-0.011	0.035	0.170	-1.172
	(0.927)	(0.855)	(0.152)	(0.105)	(0.664)	(0.922)	(0.942)	(0.247)	(0.194)	(0.817)	(0.932)	(0.857)	(0.151)	(0.664)
$\Delta GDPDEF_{t-1}$	-0.188**	0.424***	0.167**	-0.052	-1.449	-0.141	0.320**	0.092	0.075	0.576	-0.197**	0.433***	0.173**	-0.693
	(0.026)	(0.000)	(0.027)	(0.445)	(0.362)	(0.136)	(0.021)	(0.267)	(0.329)	(0.764)	(0.018)	(0.001)	(0.020)	(0.679)
TR_{t-1}	0.016	-0.116**	0.995***	0.006	0.079	0.058	-0.211**	0.926***	0.013	2.122*	0.006	-0.106*	1.001***	0.940
	(0.687)	(0.049)	(0.000)	(0.847)	(0.916)	(0.308)	(0.013)	(0.000)	(0.785)	(0.071)	(0.874)	(0.052)	(0.000)	(0.212)
$TERM_{t-1}$	0.036	-0.036	-0.023	0.961***	-3.086***	-	-	-	-	-	-	-	-	-
	(0.480)	(0.632)	(0.617)	(0.000)	(0.002)	-	-	-	-	-	-	-	-	-
LT_{t-1}	-	-	-	-	-	-0.082	0.166*	0.119*	0.918***	-1.867	-	-	-	-
	-	-	-	-	-	(0.232)	(0.099)	(0.051)	(0.000)	(0.186)	-	-	-	-
LS_{t-1}	-0.011***	0.001	-0.005*	0.009***	0.823***	-0.010***	-0.000	-0.005**	0.000	0.841***	-0.011***	0.001	-0.005*	0.831***
	(0.001)	(0.891)	(0.076)	(0.001)	(0.000)	(0.001)	(0.958)	(0.042)	(0.876)	(0.000)	(0.001)	(0.874)	(0.078)	(0.000)

Note: Numbers in bracket indicate the p-values for the t-stats. *,** and *** indicate significance at the 1%, 5% and 10% levels respectively.

Table 5.5: Coefficient estimates and t-stats for the lending standards equation

The table presents the coefficient estimates for regressions of lending standards on a set of regressors including macroeconomic variables, a measure of the stance of monetary policy and the term spread. Lending standards (LS) is the net percentage of banks reporting tightening lending standards over the last quarter (Senior Loan Officer Opinion Survey). ΔGDP is the quarterly growth rate of real GDP. $\Delta GDPDEF$ is the quarterly change in the GDP deflator. FF is the Effective Fed Funds rate. TR are Taylor rule residuals computed as the difference between the Fed Funds rate and the Taylor rule rate (see Table 5.2 for details of calculations). $TERM$ is the difference between the 10-year Treasury bond rate and the 3-month-Tbill rate. In addition each model includes in turn a variable controling for banks' profitability, measures of risk on both the liability side (core liabilities) and the asset side (securitization) and other lending terms (loan volumes and intermediation spreads). NIM is the net interst margin of commercial banks. ΔBE is the growth rate of a measure of the current profitability (or equity) of commercial banks. ΔCL is the growth rate of the (real) sum of repos and large time deposits on the liability side of commercial banks' balance sheets. ΔSEC is the growth rate of the (real) volume of asset-backed commercial paper (not including those issued by commercial banks). ΔCIL is the growth rate of the (real) volume of business loans at commercial banks. $SPREAD$ is the net percentage of banks reporting increasing spreads of loan rates over cost of funds.

	LS		LS		LS		LS		LS		LS	
ΔGDP_{t-1}	-1.122	-1.388	-0.896	-0.855	-1.221	-1.138	-1.069	-1.029	-0.809	-0.875	-1.471	-1.390
	(0.657)	(0.590)	(0.723)	(0.739)	(0.629)	(0.655)	(0.684)	(0.677)	(0.752)	(0.736)	(0.553)	(0.579)
$\Delta GDPDEF_{t-1}$	-1.586	-1.414	-1.504	-1.287	23.237	14.246	-1.685	-1.440	-1.686	-1.443	-1.560	-1.302
	(0.351)	(0.376)	(0.309)	(0.426)	(0.615)	(0.756)	(0.261)	(0.370)	(0.256)	(0.367)	(0.282)	(0.408)
FF_{t-1}	0.561	-	0.640	-	0.697	-	0.747	-	0.738	-	0.678	-
	(0.640)	-	(0.443)	-	(0.422)	-	(0.431)	-	(0.399)	-	(0.406)	-
TR_{t-1}	-	-0.493	-	0.179	-	0.141	-	0.110	-	0.150	-	0.215
	-	(0.676)	-	(0.816)	-	(0.856)	-	(0.897)	-	(0.846)	-	(0.774)
$TERM_{t-1}$	-2.578	-3.824**	-2.567**	-3.097***	-2.251	-2.932***	-2.521**	-3.095***	-3.253**	-3.586**	-3.096**	-3.634***
	(0.190)	(0.014)	(0.037)	(0.002)	(0.102)	(0.008)	(0.042)	(0.003)	(0.046)	(0.025)	(0.014)	(0.001)
NIM_{t-1}	-0.017	3.561	-	-	-	-	-	-	-	-	-	-
	(0.997)	(0.528)	-	-	-	-	-	-	-	-	-	-
ΔBE_{t-1}	-	-	0.129	0.118	-	-	-	-	-	-	-	
	-	-	(0.483)	(0.527)	-	-	-	-	-	-	-	-
ΔCL_{t-1}	-	-	-	-	24.877	15.671	-	-	-	-	-	-
	-	-	-	-	(0.591)	(0.731)	-	-	-	-	-	-
ΔSEC_{t-1}	-	-	-	-	-	-	-0.068	-0.013	-	-	-	-
	-	-	-	-	-	-	(0.682)	(0.934)	-	-	-	-
ΔCIL_{t-1}	-	-	-	-	-	-	-	-	-0.164	-0.104	-	-
	-	-	-	-	-	-	-	-	(0.526)	(0.681)	-	-
$SPREAD_{t-1}$	-	-	-	-	-	-	-	-	-	-	0.100	0.097
	-	-	-	-	-	-	-	-	-	-	(0.105)	(0.117)

Note: The equation is estimated by OLS and also includes a constant term and one lag of the dependent variable (not reported) over the period 1990Q2-2007Q4 (70 obs). Numbers in bracket indicate the p-values for the t-stats. *,** and *** indicate significance at the 1%, 5% and 10% levels respectively.

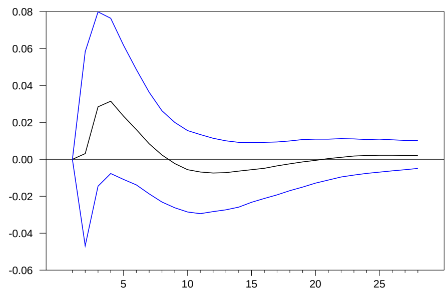
Table 5.6: Lending standards and demand for loans

The table presents coefficient estimates of a VAR estimated with 1 lag and a constant term (not reported) over the period 1991Q4-2007Q4 (65 obs). ΔGDP is the quarterly growth rate of real GDP. $\Delta GDPDEF$ is the quarterly change in the GDP deflator. FF is the Effective Fed Funds rate. TR are Taylor rule residuals computed as the difference between the Fed Funds rate and the Taylor rule rate (see Table 5.2 for details of calculations). $TERM$ is the difference between the 10-year Treasury bond rate and the 3-month Tbill rate. $DEMAND$ and LS are the net percentage of banks reporting a stronger demand for bank loans and tightening lending standards for loans over the last quarter respectively (SLOOS).

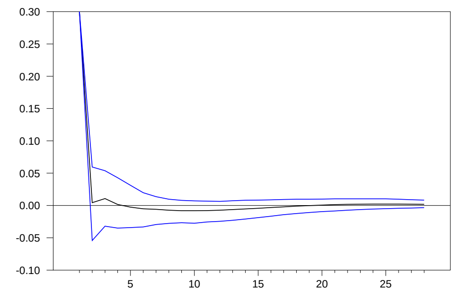
	ΔGDP	$\Delta GDPDEF$	FF	$TERM$	$DEMAND$	LS
ΔGDP_{t-1}	-0.053	0.183	0.195**	-0.098	10.480***	-0.650
	(0.707)	(0.381)	(0.038)	(0.401)	(0.006)	(0.800)
$\Delta GDPDEF_{t-1}$	-0.178*	0.361***	0.071	-0.035	-0.866	0.589
	(0.057)	(0.009)	(0.237)	(0.642)	(0.718)	(0.723)
FF_{t-1}	0.016	-0.194**	0.930***	-0.007	-0.048	2.268**
	(0.785)	(0.027)	(0.000)	(0.881)	(0.975)	(0.036)
$TERM_{t-1}$	0.038	-0.173	-0.024	0.911***	2.080	-1.535
	(0.785)	(0.117)	(0.620)	(0.000)	(0.285)	(0.256)
$DEMAND_{t-1}$	-0.001	0.002	0.008***	-0.007**	0.578***	-0.179***
	(0.740)	(0.681)	(0.001)	(0.017)	(0.000)	(0.010)
LS_{t-1}	-0.010**	-0.003	-0.003	0.001	-0.411***	0.680***
	(0.037)	(0.703)	(0.297)	(0.713)	(0.002)	(0.000)

Note: Numbers in bracket indicate the p-values for the t-stats. *,** and *** indicate significance at the 1%, 5% and 10% levels respectively.

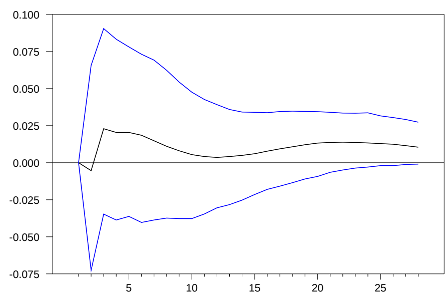
Figure 5-3: Impulse responses to a monetary shock



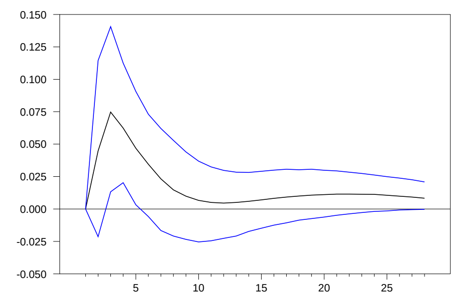
(a) Real GDP growth



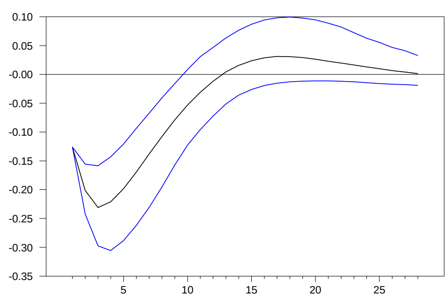
(b) Real GDP growth (LR restriction)



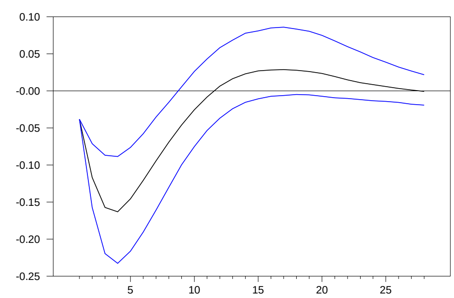
(c) Inflation



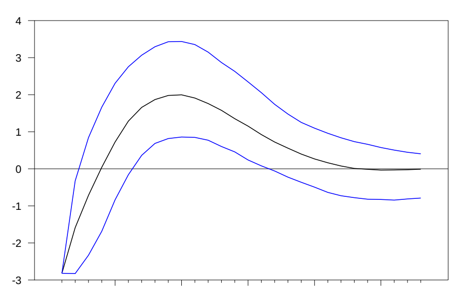
(d) Inflation (LR restriction)



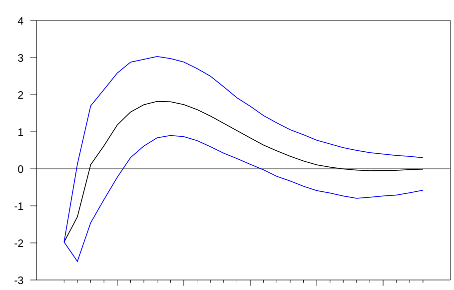
(e) Term spread



(f) Term spread (LR restriction)



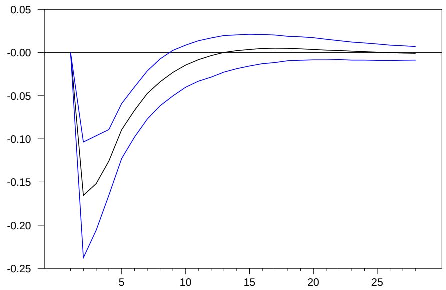
(g) Lending standards



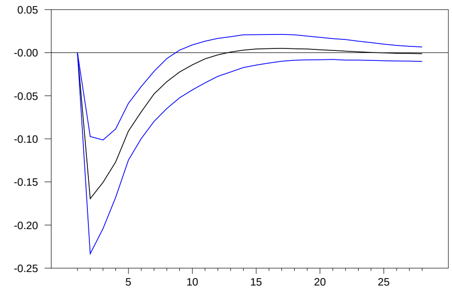
(h) Lending standards (LR restriction)

Note: We estimate a 5-variable VAR including two lags of the growth rate of real GDP, the growth rate of the GDP deflator, the Fed Funds rate, the term spread and lending standards over the period 1990Q2-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. The left-hand panels represent the impulse responses obtained using a Cholesky decomposition. The right-hand panels represents the impulse responses obtained by placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of GDP growth to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity). Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

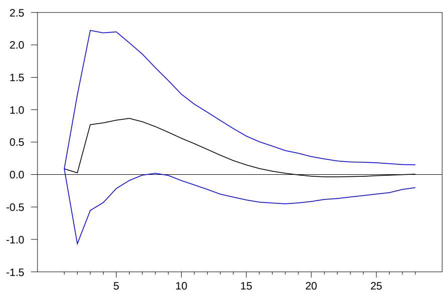
Figure 5-4: Lending standards



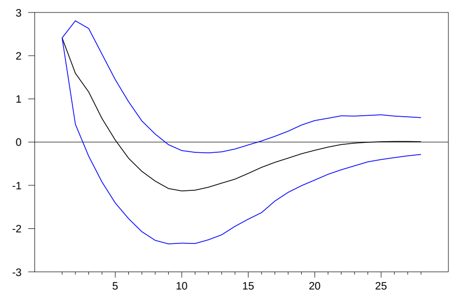
(a) LS → GDP growth



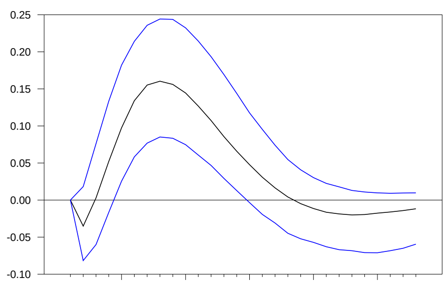
(b) LS → GDP growth (LR restriction)



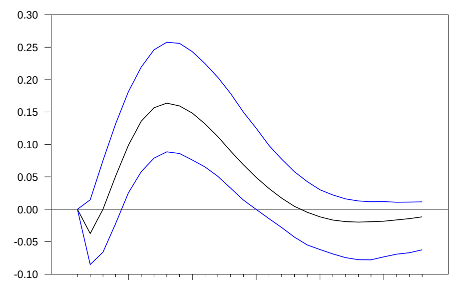
(c) GDP growth → LS



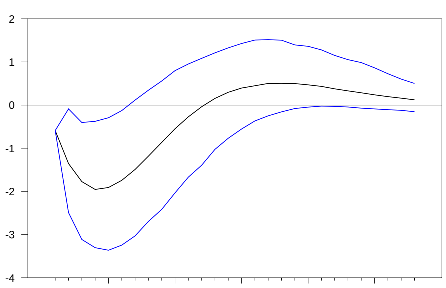
(d) GDP growth → LS (LR restriction)



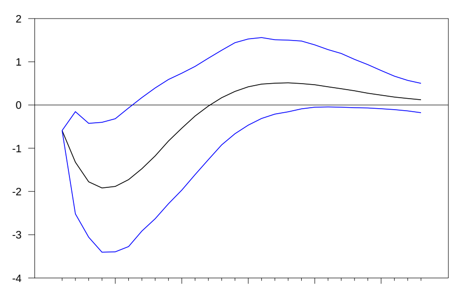
(e) LS → Term spread



(f) LS → Term spread (LR restriction)



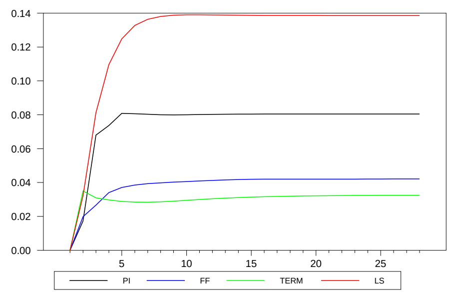
(g) Term spread → LS



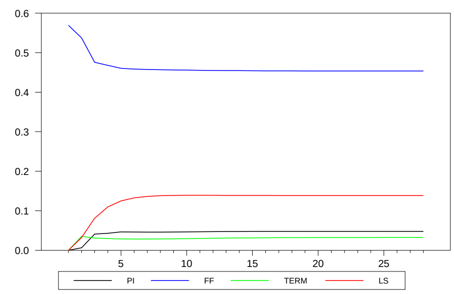
(h) Term spread → LS (LR restriction)

Note: We estimate a 5-variable VAR including two lags of the growth rate of real GDP, the growth rate of the GDP deflator, the Fed Funds rate, the term spread and lending standards over the period 1990Q2-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. The left-hand panels represent the impulse responses obtained using a Cholesky decomposition. The right-hand panels represents the impulse responses obtainedby placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of GDP growth to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity). Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

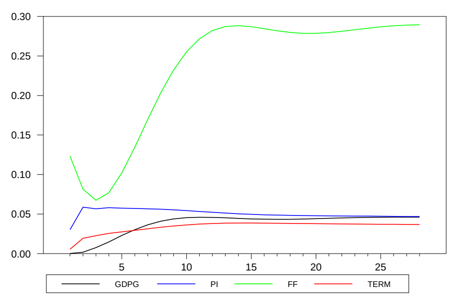
Figure 5-5: Forecast error variance decomposition of real GDP growth and lending standards



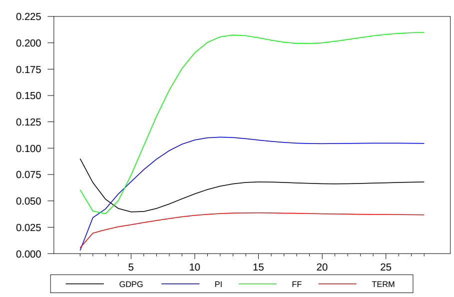
(a) Real GDP growth



(b) Real GDP growth (LR restriction)



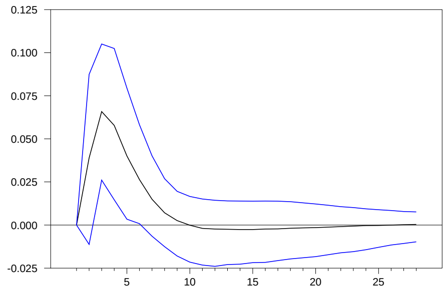
(c) Lending standards



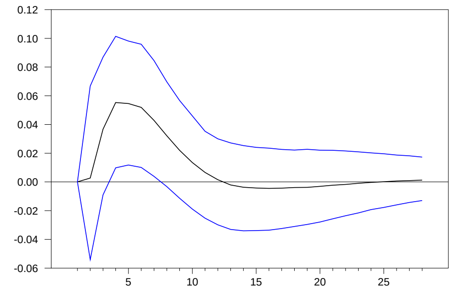
(d) Lending standards (LR restriction)

Note: We estimate a 5-variable VAR including two lags of the growth rate of real GDP, the growth rate of the GDP deflator, the Fed Funds rate, the term spread and lending standards over the period 1990Q2-2007Q4. Forecast error variance decompostions (FEVD) represents the fraction of the variance of the expected forecast error that can be attributed to shocks to the other variables. The left-hand panels represent the FEVD obtained using a Cholesky decomposition. The right-hand panels represents the FEVD obtainedby placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of GDP growth to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity).

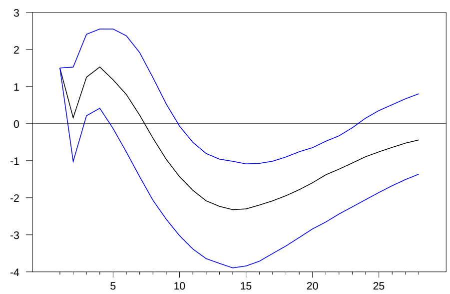
Figure 5-6: Impulse responses to a monetary shock



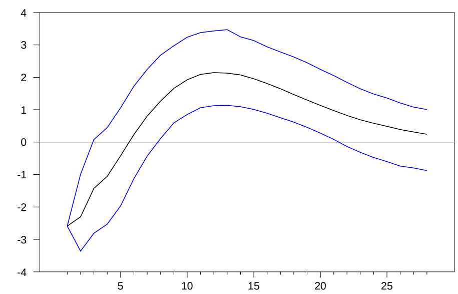
(a) Real GDP



(b) Inflation



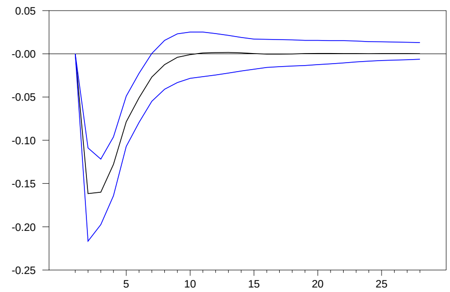
(c) Loan demand



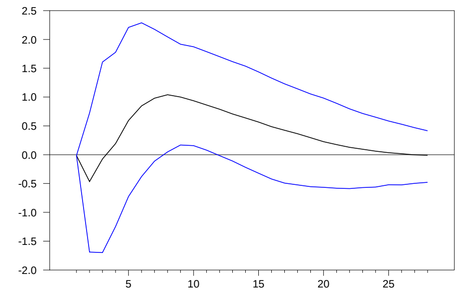
(d) Lending standards

Note: We estimate a 5-variable VAR including 2 lags of the growth rate of real GDP, the growth rate of the GDP deflator, the Fed Funds rate, demand for loans and lending standards (in that order) over the period 1991Q4-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

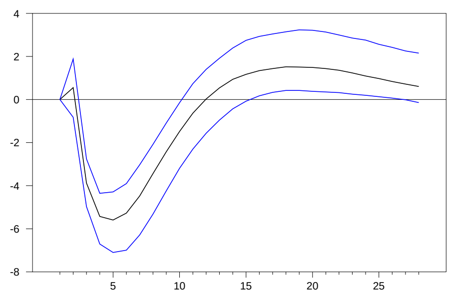
Figure 5-7: Lending standards



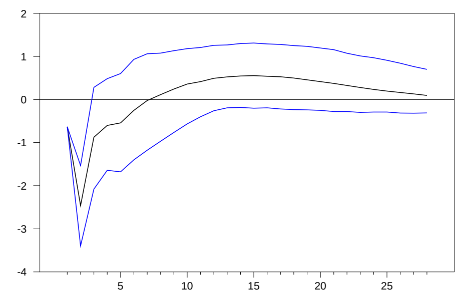
(a) LS → GDP



(b) GDP → LS



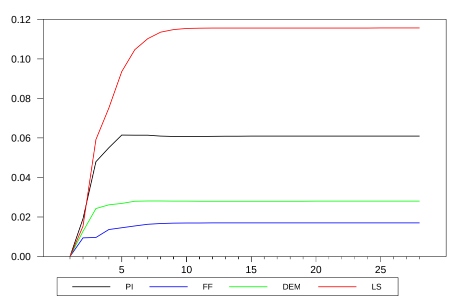
(c) LS → Loan demand



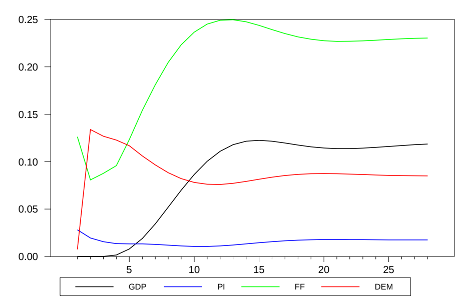
(d) Loan demand → LS

Note: We estimate a 5-variable VAR including 2 lags of the growth rate of real GDP, the growth rate of the GDP deflator, the Fed Funds rate, demand for loans and lending standards (in that order) over the period 1991Q4-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

Figure 5-8: Forecast error variance decomposition of real GDP and lending standards



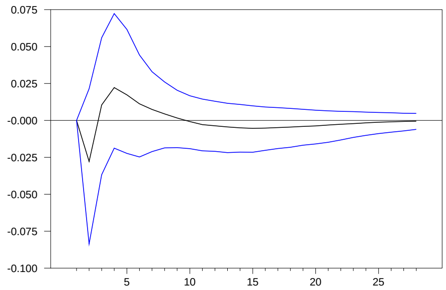
(a) GDP



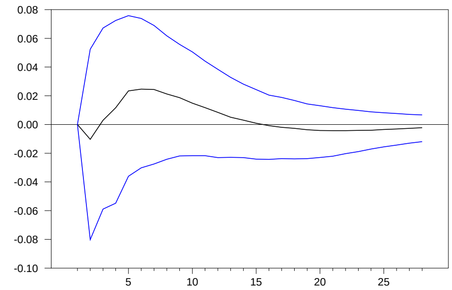
(b) Lending standards

Note: We estimate a 5-variable VAR including 2 lags of the growth rate of real GDP, the growth rate of the GDP deflator, the Fed Funds rate, demand for loans and lending standards (in that order) over the period 1991Q4-2007Q4. Forecast error variance decompostions (FEVD) represents the fraction of the variance of the expected forecast error that can be attributed to shocks to the other variables. Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

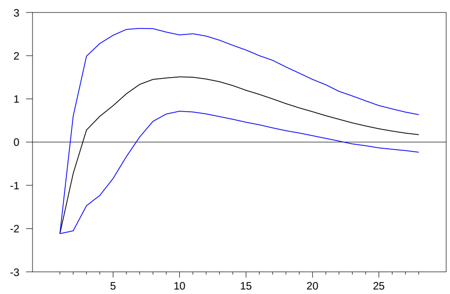
Figure 5-9: Impulse responses to a monetary shock



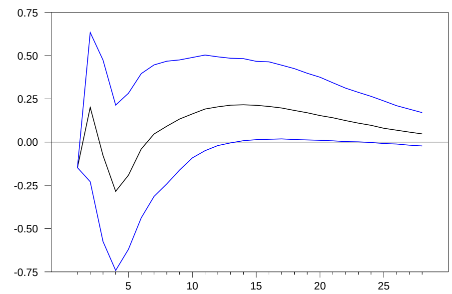
(a) Real GDP



(b) Inflation



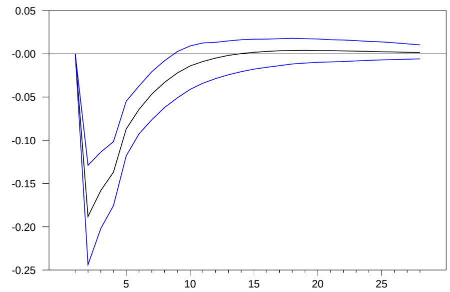
(c) Lending standards



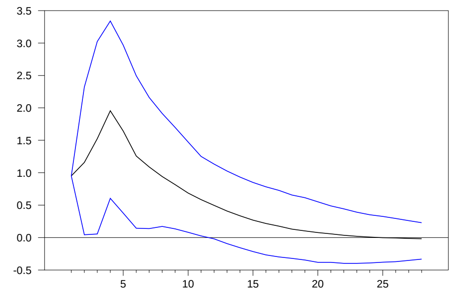
(d) VIX

Note: We estimate a 5-variable VAR including 2 lags of the growth rate of real GDP, the growth rate of the GDP deflator, the Fed Funds rate, lending standards and the VIX index (in that order) over the period 1990Q2-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

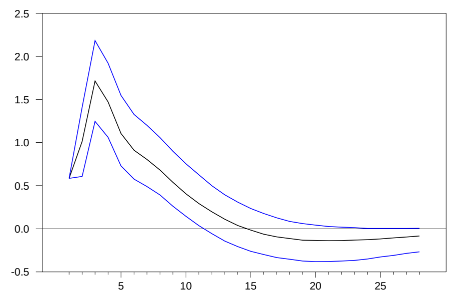
Figure 5-10: Lending standards



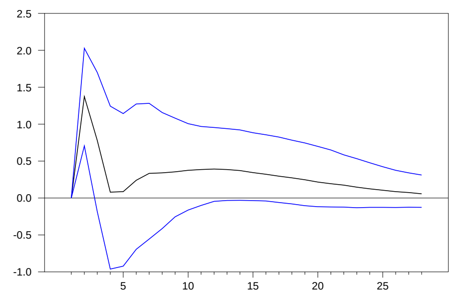
(a) LS → GDP



(b) GDP → LS



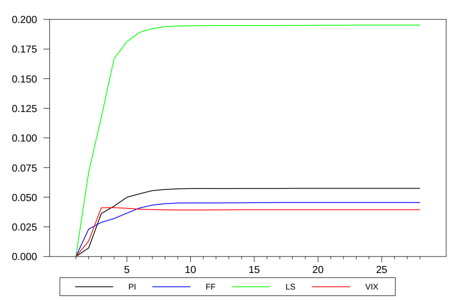
(c) LS → VIX



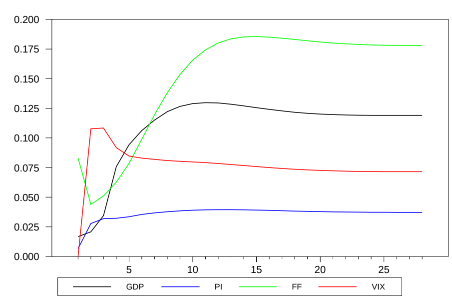
(d) VIX → LS

Note: We estimate a 5-variable VAR including 2 lags of the growth rate of real GDP, the growth rate of the GDP deflator, the Fed Funds rate, lending standards and the VIX index (in that order) over the period 1990Q2-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

Figure 5-11: Forecast error variance decomposition of real GDP and lending standards



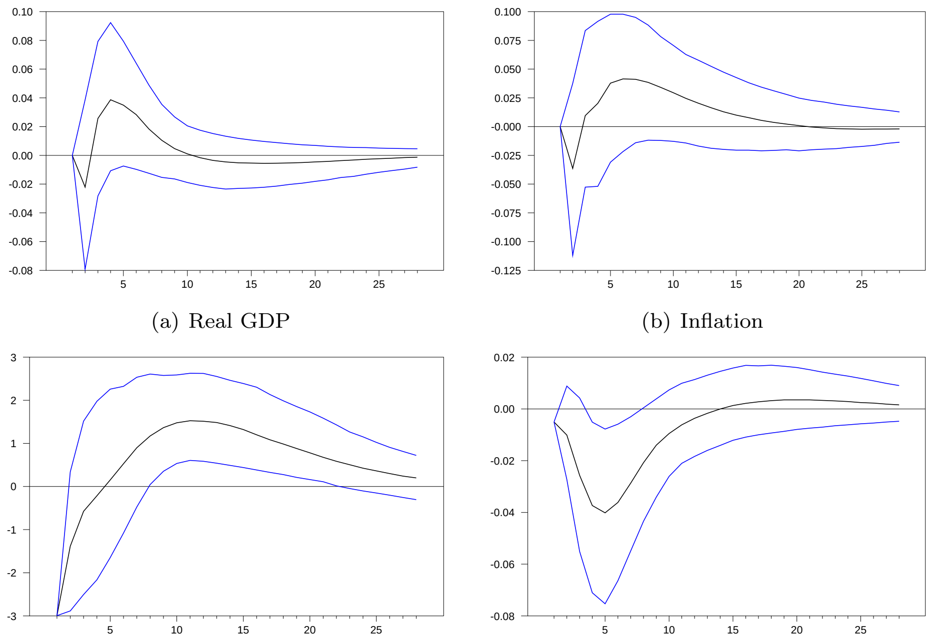
(a) GDP



(b) Lending standards

Note: We estimate a 5-variable VAR including 2 lags of the growth rate of real GDP, the growth rate of the GDP deflator, the Fed Funds rate, lending standards and the VIX index (in that order) over the period 1990Q2-2007Q4. Forecast error variance decompostions (FEVD) represents the fraction of the variance of the expected forecast error that can be attributed to shocks to the other variables. Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

Figure 5-12: Impulse responses to a monetary shock



(a) Real GDP

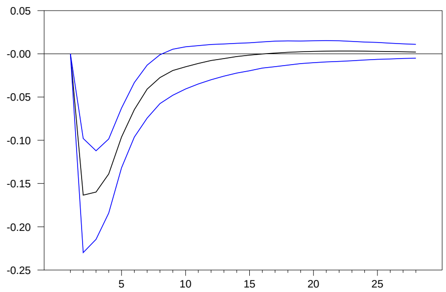
(b) Inflation

(c) Lending standards

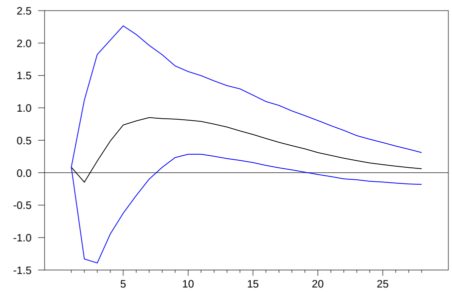
(d) BAA-AAA spread

Note: We estimate a 5-variable VAR including 2 lags of the growth rate of real GDP, the growth rate of the GDP deflator, the Fed Funds rate, lending standards and the BAA-AAA spread (in that order) over the period 1990Q2-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

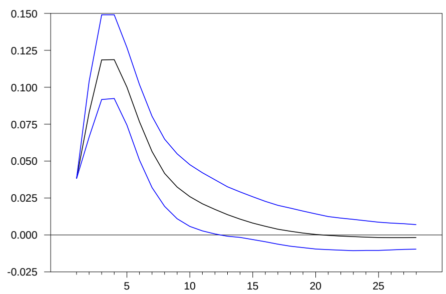
Figure 5-13: Lending standards



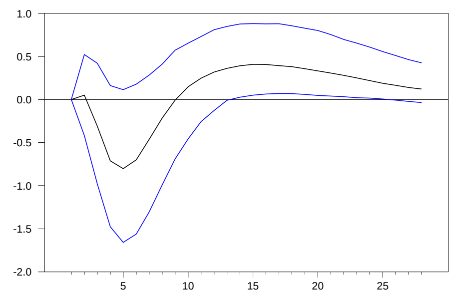
(a) LS → GDP



(b) GDP → LS



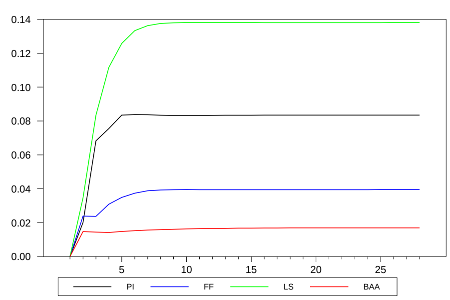
(c) LS → BAA-AAA spread



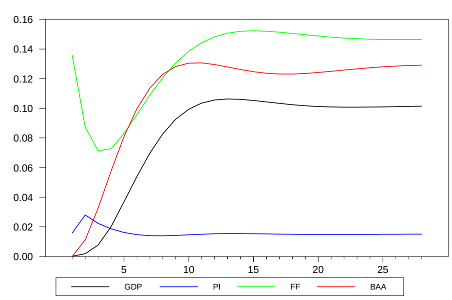
(d) BAA-AAA spread → LS

Note: We estimate a 5-variable VAR including 2 lags of the growth rate of real GDP, the growth rate of the GDP deflator, the Fed Funds rate, lending standards and the BAA-AAA spread (in that order) over the period 1990Q2-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

Figure 5-14: Forecast error variance decomposition of real GDP and lending standards



(a) GDP



(b) Lending standards

Note: We estimate a 5-variable VAR including 2 lags of the growth rate of real GDP, the growth rate of the GDP deflator, the Fed Funds rate, lending standards and the BAA-AAA spread (in that order) over the period 1990Q2-2007Q4. Forecast error variance decompostions (FEVD) represents the fraction of the variance of the expected forecast error that can be attributed to shocks to the other variables. Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

Appendix A: Sensitivity checks

In this appendix, we present impulse response functions and forecast error variance decompostions for alternative specifications of the model.

In the first specification, we order the lending standards variable in third position (instead of last in the baseline specification).

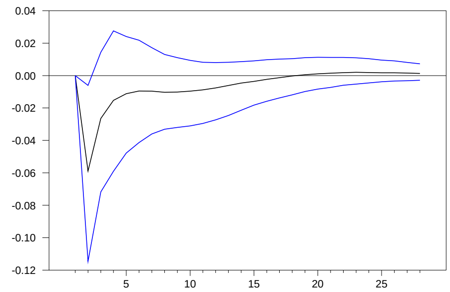
The second specification contains an output gap measure (where the log of real GDP is detrended using the HP filter and data over the period 1970Q1-2015Q1), a measure of the annual inflation rate (calculated as the fourth difference of the log of GDP deflator) and the Fed Funds rate so that it is testing more directly the predictions derived from a fully microfunded DSGE model.

In the third specification, we specify nonstationary variables in log levels, so as to preserve the information contained in the potential cointegrating relationships among the time series (Sims, Stock and Watson). In this case we only impose a Choleski decomposition to identify orthogonal innovations.

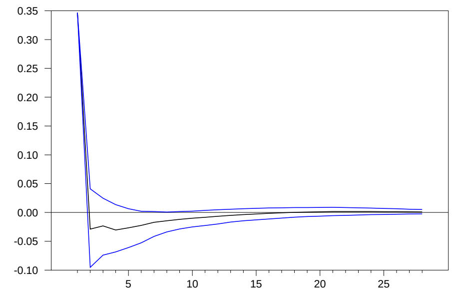
In the fourth specification, we conduct the analysis ove the full sample (1990Q2-2015Q1).

Specification 1: $Y_t = [\Delta GDP_t, \Delta GDPDEF_t, S_t, FF_t, TERM_t]$

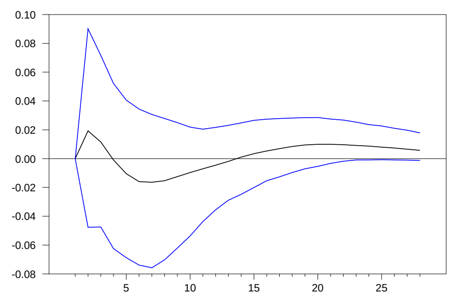
Figure 5-15: Impulse responses to a monetary shock



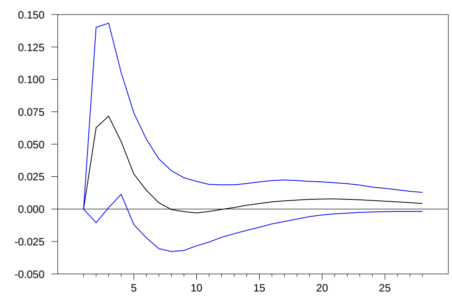
(a) Real GDP growth



(b) Real GDP growth (LR restriction)



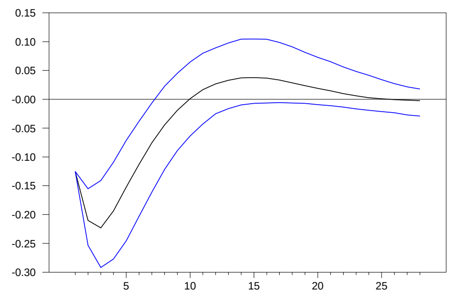
(c) Inflation



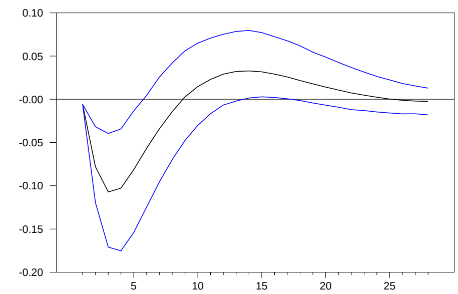
(d) Inflation (LR restriction)

Note: We estimate a 5-variable VAR including two lags of the growth rate of real GDP, the growth rate of the GDP deflator, lending standards, the Fed Funds rate and the term spread (in that order) over the period 1990Q2-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. The left-hand panels represent the impulse responses obtained using a Cholesky decomposition. The right-hand panels represents the impulse responses obtained by placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of GDP growth to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity). Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

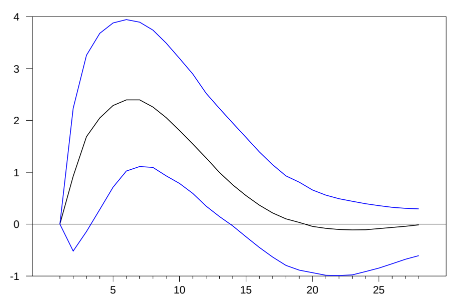
Figure 5-16: Impulse responses to a monetary shock (continued)



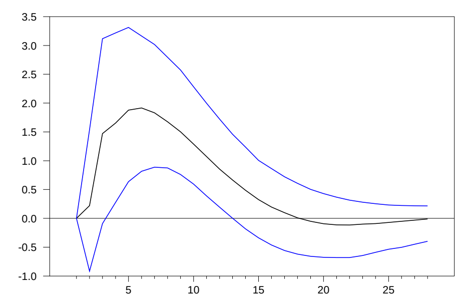
(a) Term spread



(b) Term spread (LR restriction)



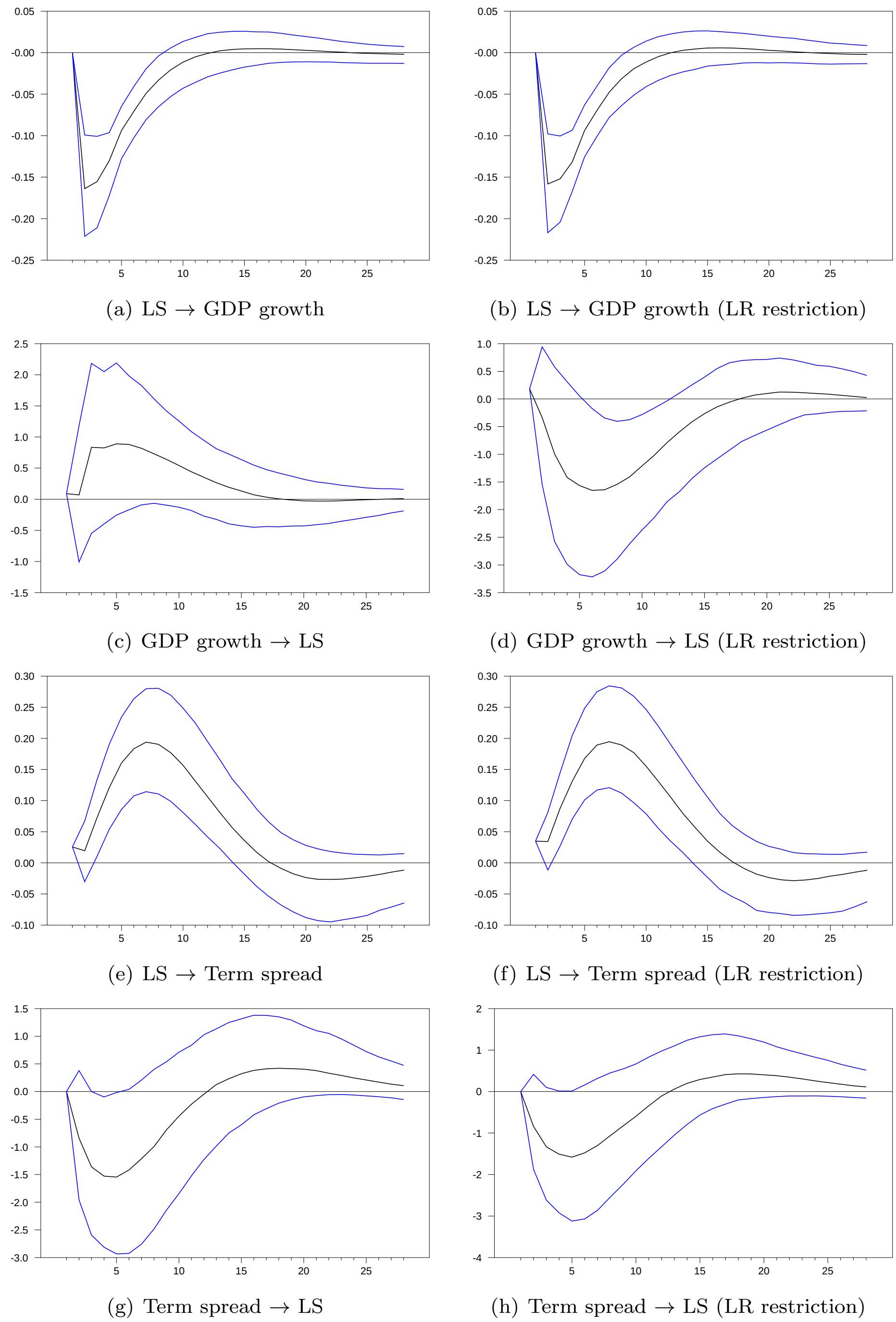
(c) Lending standards



(d) Lending standards (LR restriction)

Note: We estimate a 5-variable VAR including two lags of the growth rate of real GDP, the growth rate of the GDP deflator, lending standards, the Fed Funds rate and the term spread (in that order) over the period 1990Q2-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. The left-hand panels represent the impulse responses obtained using a Cholesky decomposition. The right-hand panels represents the impulse responses obtained by placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of GDP growth to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity). Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

Figure 5-17: Lending standards



(a) LS → GDP growth

(b) LS → GDP growth (LR restriction)

(c) GDP growth → LS

(d) GDP growth → LS (LR restriction)

(e) LS → Term spread

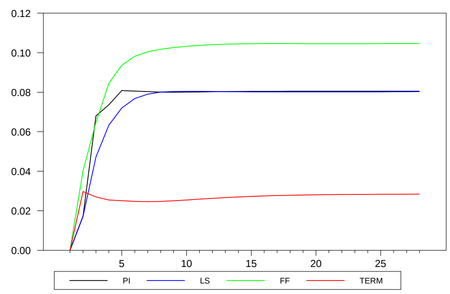
(f) LS → Term spread (LR restriction)

(g) Term spread → LS

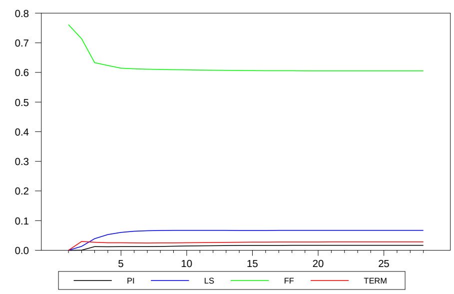
(h) Term spread → LS (LR restriction)

Note: We estimate a 5-variable VAR including two lags of the growth rate of real GDP, the growth rate of the GDP deflator, lending standards, the Fed Funds rate and the term spread (in that order) over the period 1990Q2-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. The left-hand panels represent the impulse responses obtained using a Cholesky decomposition. The right-hand panels represents the impulse responses obtained by placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of GDP growth to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity). Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

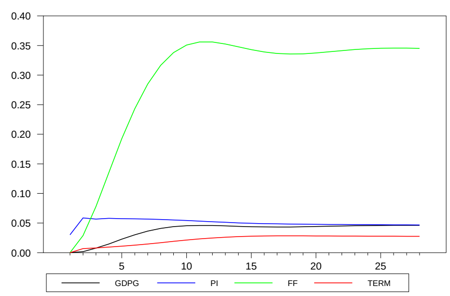
Figure 5-18: Forecast error variance decomposition of real GDP growth and lending standards



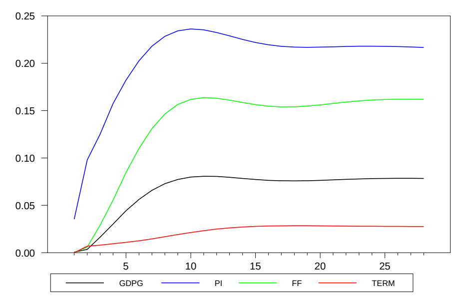
(a) Real GDP growth



(b) Real GDP growth (LR restriction)



(c) Lending standards

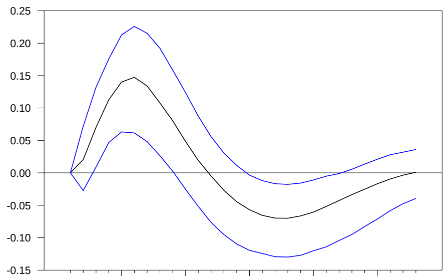


(d) Lending standards (LR restriction)

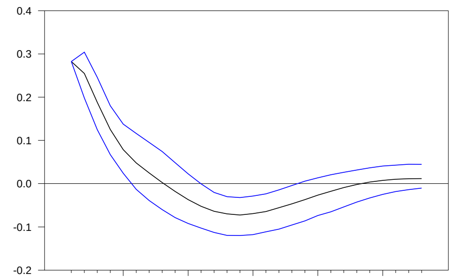
Note: We estimate a 5-variable VAR including two lags of the growth rate of real GDP, the growth rate of the GDP deflator, lending standards, the Fed Funds rate and the term spread over the period 1990Q2-2007Q4. Forecast error variance decompostions (FEVD) represents the fraction of the variance of the expected forecast error that can be attributed to shocks to the other variables. The left-hand panels represent the FEVD obtained using a Cholesky decomposition. The right-hand panels represents the FEVD obtained by placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of GDP growth to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity).

Specification 2: $Y_t = [YGAP_t, \Delta GDPDEF_t, FF_t, TERM_t, S_t]$

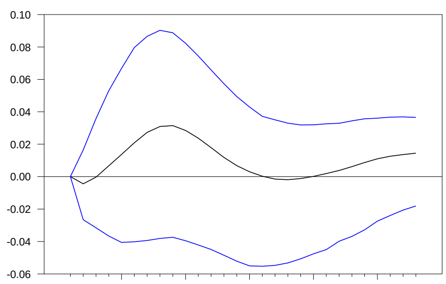
Figure 5-19: Impulse responses to a monetary shock



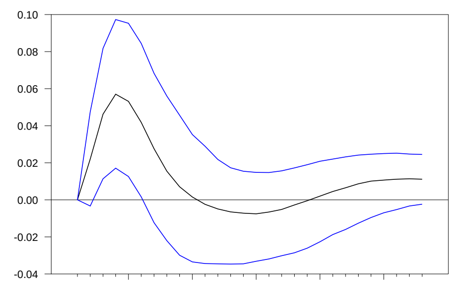
(a) Output gap



(b) Output gap (LR restriction)



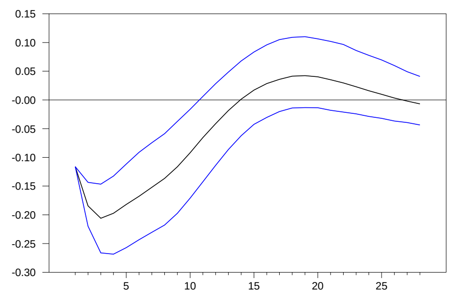
(c) Inflation



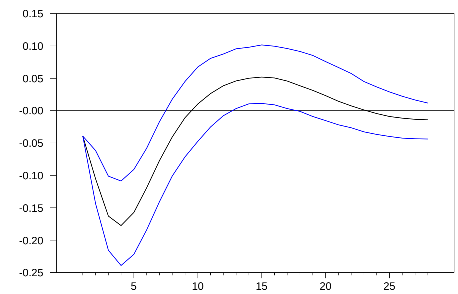
(d) Inflation (LR restriction)

Note: We estimate a 5-variable VAR including 2 lags of the output gap, the annual growth rate of the GDP deflator, the Fed Funds rate, the term spread and lending standards (in that order) over the period 1990Q2-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. The left-hand panels represent the impulse responses obtained using a Cholesky decomposition. The right-hand panels represents the impulse responses obtained by placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of output gap to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity). Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

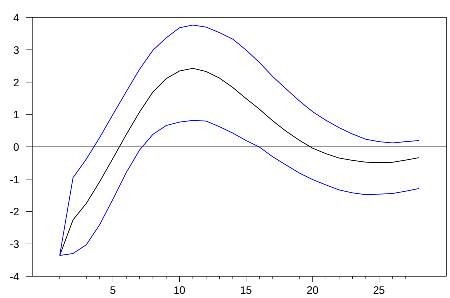
Figure 5-20: Impulse responses to a monetary shock (continued)



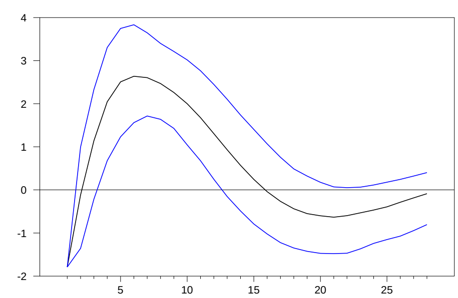
(a) Term spread



(b) Term spread (LR restriction)



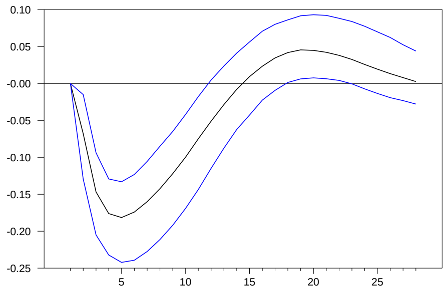
(c) Lending standards



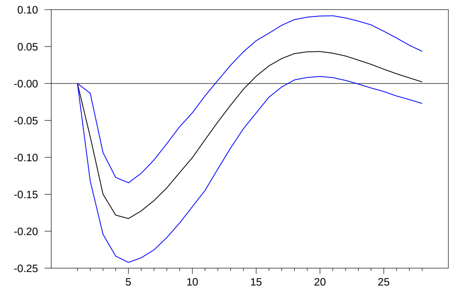
(d) Lending standards (LR restriction)

Note: We estimate a 5-variable VAR including 2 lags of the output gap, the annual growth rate of the GDP deflator, the Fed Funds rate, the term spread and lending standards (in that order) over the period 1990Q2-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. The left-hand panels represent the impulse responses obtained using a Cholesky decomposition. The right-hand panels represents the impulse responses obtained by placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of output gap to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity). Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

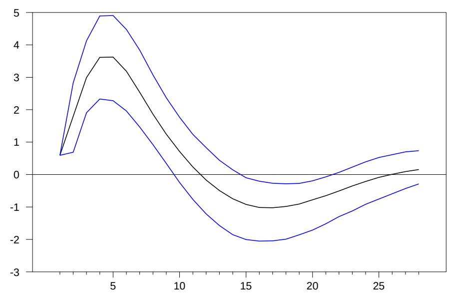
Figure 5-21: Lending standards



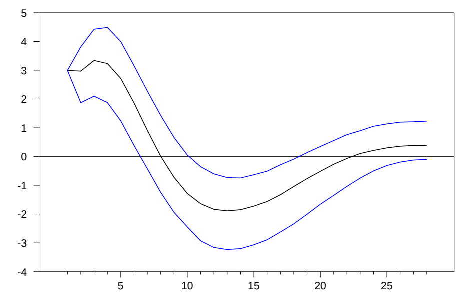
(a) LS → Output gap



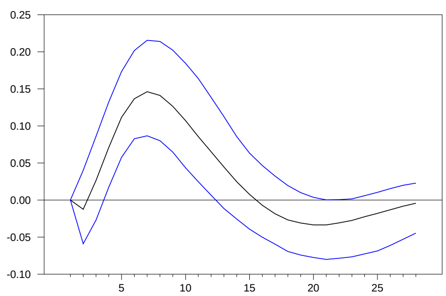
(b) LS → Output gap (LR restriction)



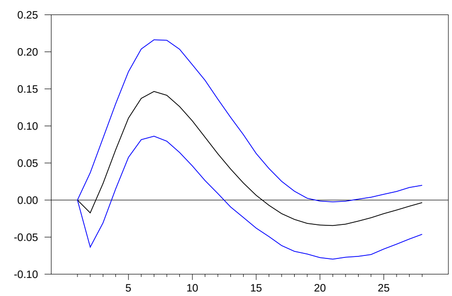
(c) Output gap → LS



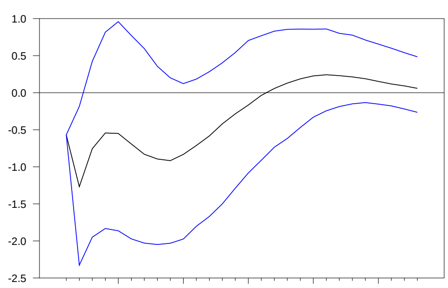
(d) Output gap → LS (LR restriction)



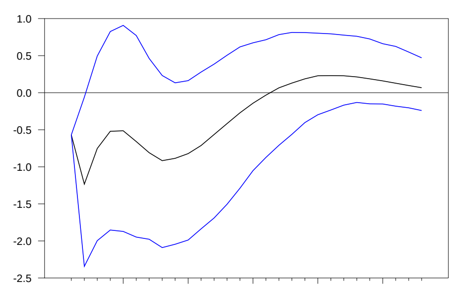
(e) LS → Term spread



(f) LS → Term spread (LR restriction)



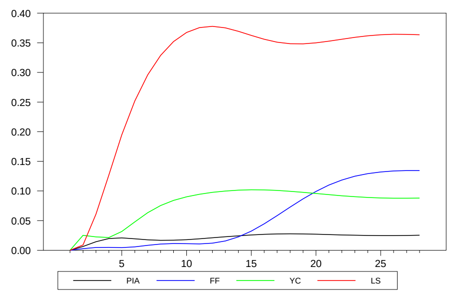
(g) Term spread → LS



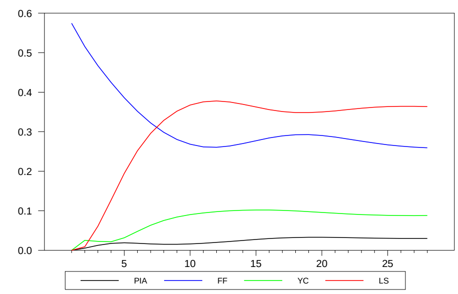
(h) Term spread → LS (LR restriction)

Note: We estimate a 5-variable VAR including 2 lags of the output gap, the annual growth rate of the GDP deflator, the Fed Funds rate, the term spread and lending standards (in that order) over the period 1990Q2-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. The left-hand panels represent the impulse responses obtained using a Cholesky decomposition. The right-hand panels represents the impulse responses obtained by placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of output gap to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity). Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

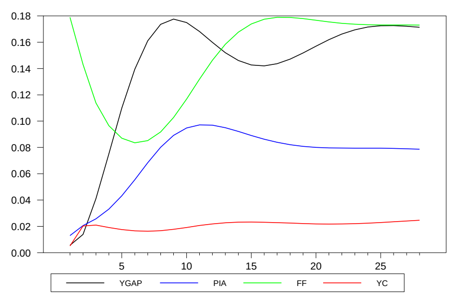
Figure 5-22: Forecast error variance decomposition of real GDP growth and lending standards



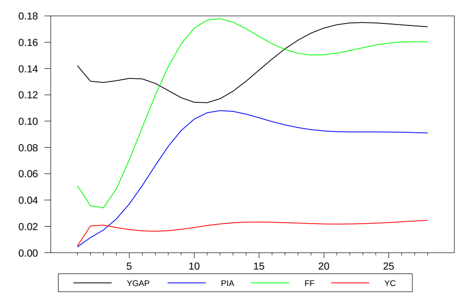
(a) Output gap



(b) Output gap (LR restriction)



(c) Lending standards

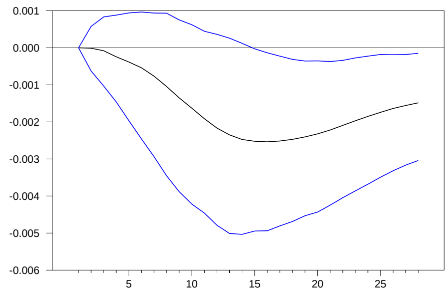


(d) Lending standards (LR restriction)

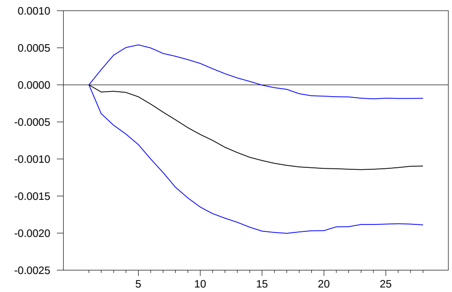
Note: We estimate a 5-variable VAR including 2 lags of the output gap, the annual growth rate of the GDP deflator, the Fed Funds rate, the term spread and lending standards over the period 1990Q2-2007Q4. Forecast error variance decompostions (FEVD) represents the fraction of the variance of the expected forecast error that can be attributed to shocks to the other variables. The left-hand panels represent the FEVD obtained using a Cholesky decomposition. The right-hand panels represents the FEVD obtainedby placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of output gap to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity).

Specification 3: $Y_t = [GDP_t, GDPDEF_t, FF_t, TERM_t, S_t]$

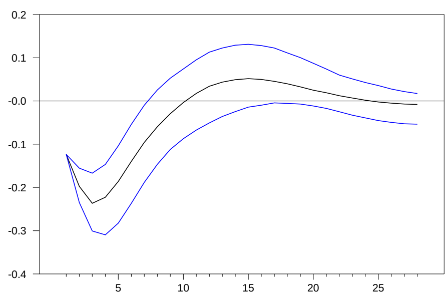
Figure 5-23: Impulse responses to a monetary shock



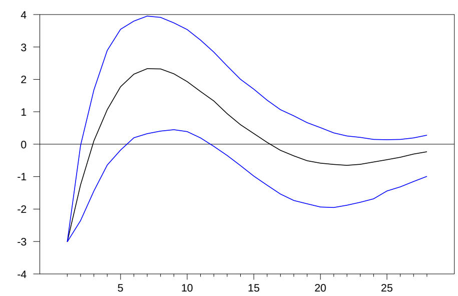
(a) Real GDP



(b) Inflation



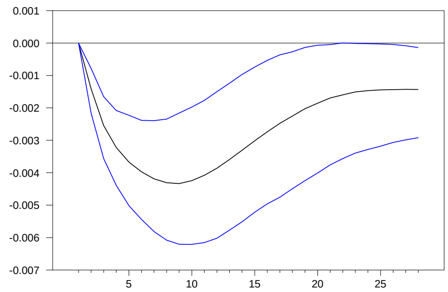
(c) Term spread



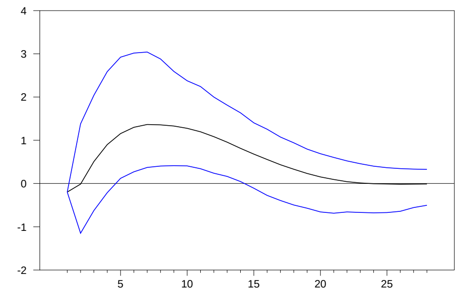
(d) Lending standards

Note: We estimate a 5-variable VAR including 2 lags of the log of real GDP, the log of the GDP deflator, the Fed Funds rate, the term spread and lending standards (in that order) over the period 1990Q2-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

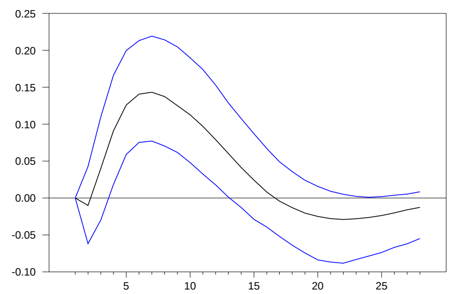
Figure 5-24: Lending standards



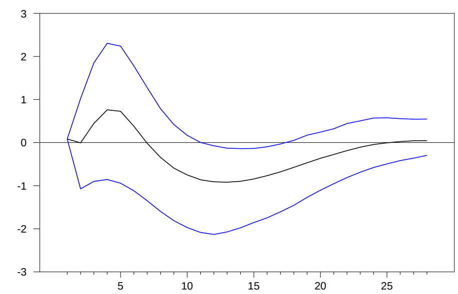
(a) LS → GDP



(b) GDP → LS



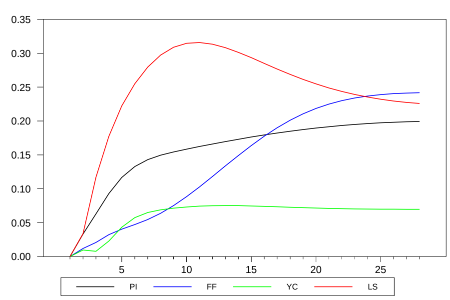
(c) LS → Term spread



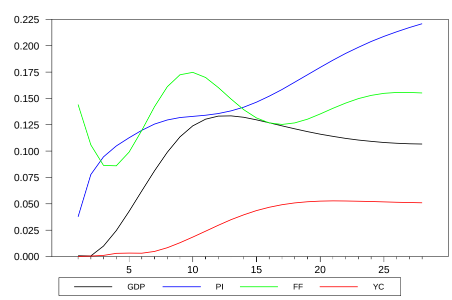
(d) Term spread → LS

Note: We estimate a 5-variable VAR including 2 lags of the log of real GDP, the log of the GDP deflator, the Fed Funds rate, the term spread and lending standards (in that order) over the period 1990Q2-2007Q4. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

Figure 5-25: Forecast error variance decomposition of real GDP and lending standards



(a) GDP

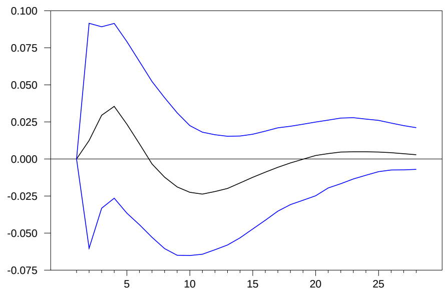


(b) Lending standards

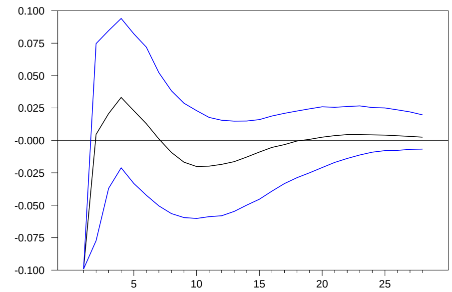
Note: We estimate a 5-variable VAR including 2 lags of the log of real GDP, the log of the GDP deflator, the Fed Funds rate, the term spread and lending standards (in that order) over the period 1990Q2-2007Q4. Forecast error variance decompostions (FEVD) represents the fraction of the variance of the expected forecast error that can be attributed to shocks to the other variables. Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

Specification 4: $Y_t = [\Delta GDP_t, \Delta GDPDEF_t, FF_t, TERM_t, S_t]$ (full sample)

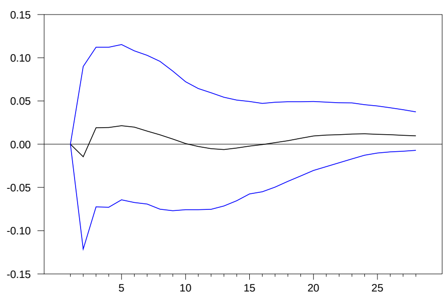
Figure 5-26: Impulse responses to a monetary shock



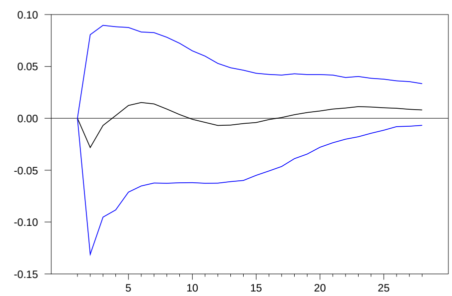
(a) Real GDP growth



(b) Real GDP growth (LR restriction)



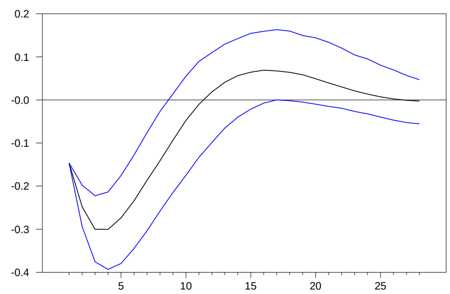
(c) Inflation



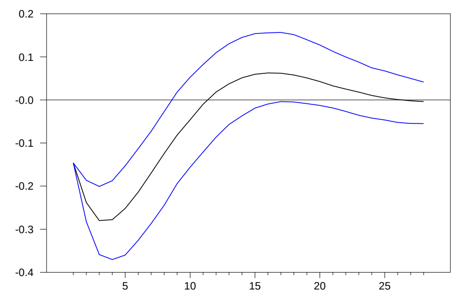
(d) Inflation (LR restriction)

Note: We estimate a 5-variable VAR including two lags of the growth rate of real GDP, the growth rate of the GDP deflator, lending standards, the Fed Funds rate and the term spread (in that order) over the period 1990Q2-2015Q1. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. The left-hand panels represent the impulse responses obtained using a Cholesky decomposition. The right-hand panels represents the impulse responses obtained by placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of GDP growth to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity). Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

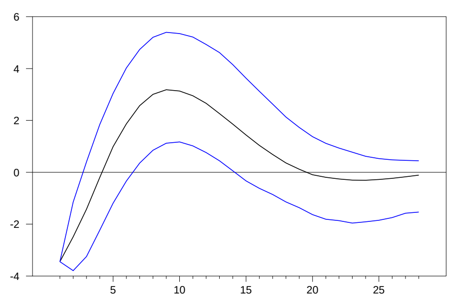
Figure 5-27: Impulse responses to a monetary shock (continued)



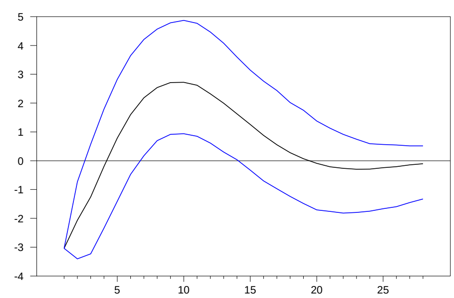
(a) Term spread



(b) Term spread (LR restriction)



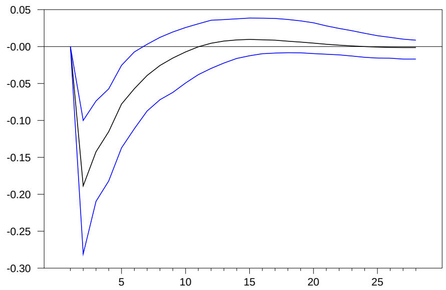
(c) Lending standards



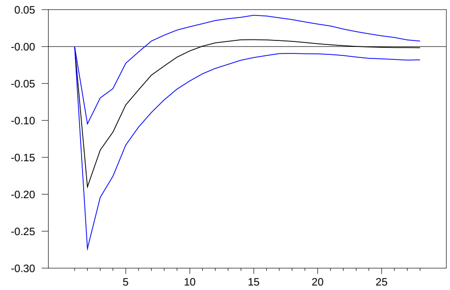
(d) Lending standards (LR restriction)

Note: We estimate a 5-variable VAR including two lags of the growth rate of real GDP, the growth rate of the GDP deflator, lending standards, the Fed Funds rate and the term spread (in that order) over the period 1990Q2-2015Q1. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. The left-hand panels represent the impulse responses obtained using a Cholesky decomposition. The right-hand panels represents the impulse responses obtained by placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of GDP growth to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity). Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

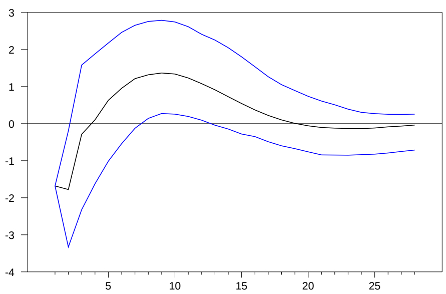
Figure 5-28: Lending standards



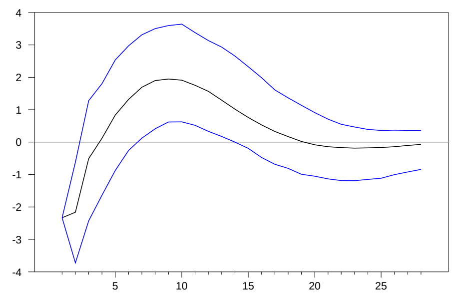
(a) LS → GDP growth



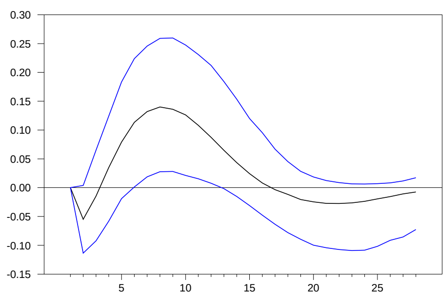
(b) LS → GDP growth (LR restriction)



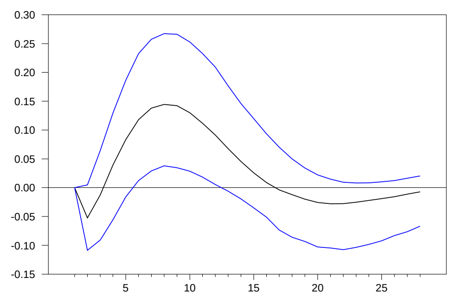
(c) GDP growth → LS



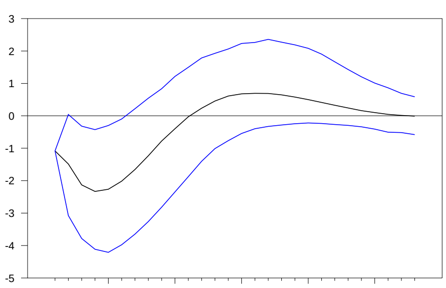
(d) GDP growth → LS (LR restriction)



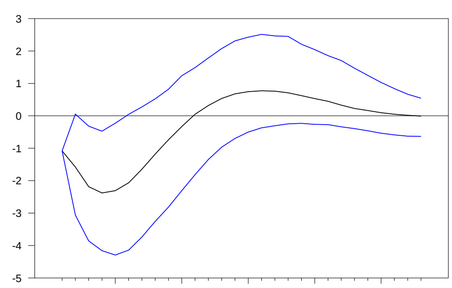
(e) LS → Term spread



(f) LS → Term spread (LR restriction)



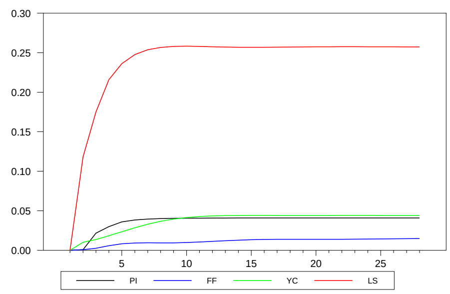
(g) Term spread → LS



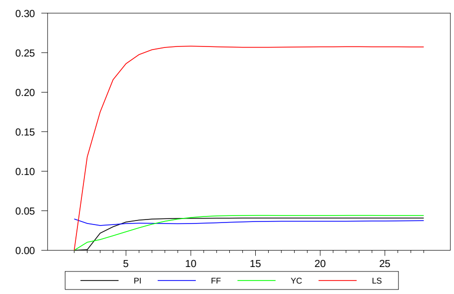
(h) Term spread → LS (LR restriction)

Note: We estimate a 5-variable VAR including two lags of the growth rate of real GDP, the growth rate of the GDP deflator, lending standards, the Fed Funds rate and the term spread (in that order) over the period 1990Q2-2015Q1. Impulse responses are to a 1 std deviation shock to the Fed Funds rate. The left-hand panels represent the impulse responses obtained using a Cholesky decomposition. The right-hand panels represents the impulse responses obtained by placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of GDP growth to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity). Error bands represent a 90% confidence interval estimating using 1000 bootstrap replications.

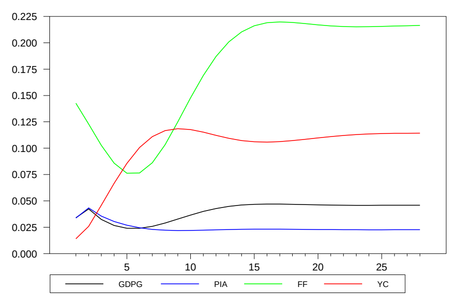
Figure 5-29: Forecast error variance decomposition of real GDP growth and lending standards



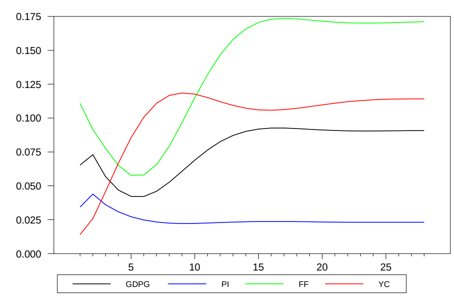
(a) Real GDP growth



(b) Real GDP growth (LR restriction)



(c) Lending standards



(d) Lending standards (LR restriction)

Note: We estimate a 5-variable VAR including two lags of the growth rate of real GDP, the growth rate of the GDP deflator, lending standards, the Fed Funds rate and the term spread over the period 1990Q2-2015Q1. Forecast error variance decompostions (FEVD) represents the fraction of the variance of the expected forecast error that can be attributed to shocks to the other variables. The left-hand panels represent the FEVD obtained using a Cholesky decomposition. The right-hand panels represents the FEVD obtained by placing the same short-run restrictions on the matrix of contemporaneous relations but replacing the short-run restriction on the contemporaneous response of GDP growth to a monetary shock by a long-run restriction (such that monetary policy has no long-run effect on real activity).

Appendix B: VAR methodology

What do impulse response functions do? Autocorrelation functions describe the timing pattern of the comovement between any two time series but do not allow discriminating between all the (potentially correlated) forces that may influence the strength, direction and timing of these relationships. Given the variables included in the VAR and the particular identification of shocks adopted, impulse responses can motivate the following questions: do we observe a significant effect of some specific shock on other variables as predicted by the theory? What are the direction, magnitude and timing of these effects? In this way we are able to follow the transmission of any one shock through the system. The more credible the identified shock in the sense that identification restrictions are strongly grounded in theory, the more explanatory power the impulse response will get because it allows a more clear-cut interpretation of the results. Hence the mapping of the theory to be tested with an appropriate identification scheme has generated an important stream of research in the VAR literature, some of these results will be discussed below.

We first describe the econometric theory behind impulse response functions. Assuming the time series in Y_t are jointly covariance stationary, the Wold theorem says that there exists a unique $MA(\infty)$ representation of this process such that X_t can be expressed as a linear combination of its current and past linear forecast errors[34]:

$$Y_t = \epsilon_t + \Psi_1 \epsilon_{t-1} + \Psi_2 \epsilon_{t-2} + \cdots$$

where Ψ_i is a 5×5 matrix and ϵ_t is a 5-dimensional white noise process (not necessarily normally distributed) representing the forecast errors of a linear regression of Y_t on $\{Y_t, Y_{t-1}, \cdots\}$. By definition, the $MA(\infty)$ polynomial is invertible and the Wold representation can be approximated by a AR(p) process and estimated using a VAR model:

$$Y_t = B(L)Y_{t-1} + \epsilon_t \quad \text{var}(\epsilon_t) = \Sigma \tag{5.2}$$

$$[I - B(L)L]Y_t = \epsilon_t \tag{5.3}$$

$$Y_t = [I - B(L)L]^{-1}\epsilon_t \tag{5.4}$$

$$Y_t = A(L)^{-1}\epsilon_t \quad A(L) = [I - B(L)L] \tag{5.5}$$

Because the $B(L)$ polynomial is invertible (the process governing Y_t is assumed to be stationary), (5.2) admits an $MA(\infty)$ representation given by (5.5), describing the set of impulse response functions of each variables to each shocks.

The Wold representation captures the second moment properties of the times series, assuming the shocks hitting the system are linearly forecastable (ie the data-generating process is well represented by a linear model). It is to the extent that these assumptions appear reasonable that the estimated impulse response functions provide a useful tool to examine the dynamics of a system of endogenous variables and shed light on the pattern of autocorrelation functions.

[34] The main reference used in this section is Cochrane (2005).

However, for this study to yield a meaningful economic interpretation, additional assumptions are required. Taking into account the dynamic relationships between all the variables in the system, the impulse response function of variable GDP_t to any shock ϵ_{it} gives the change of the subsequent realizations of the variables after a one-time unit shock to ϵ_{it}, everything else remaining equal. But since Σ may not be diagonal, each individual shock could potentially represent a mixture of different shocks and it is difficult to assign a clear interpretation to the resulting path, thus reducing the explanatory power of the procedure. Furthermore, there exists an infinity of observationally equivalent[35] combinations of shocks that can be constructed by multiplying each side of (5.5) by the same matrix.

To allow for a meaningful use of the estimated (reduced-form) system (5.2), we want to transform it into an appropriate structural form by the choice of the matrix Q:

$$Y_t = A(L)^{-1}Q\upsilon_t \quad \epsilon_t = Q\upsilon_t \tag{5.6}$$

$$Y_t = D(L)\upsilon_t \quad D(L) = A(L)^{-1}Q \tag{5.7}$$

where $D(0)$ does not necessarily equal I. The first set of identifying restrictions is the requirement that the new shocks be orthogonal to each other so that the matrix $\text{var}(\upsilon_t) = Q\Sigma Q' = I$[36]. Because there exists an infinite number of matrices Q satisfying this condition, a second set of restrictions on $D(L)$ is needed to identify the elements on Q.

Sims (1980) advocated making assumptions regarding the contemporaneous relationships between the variables, which corresponds to restricting the parameters in $D(L)$ when $L = 0$:

$$D(0) = Q \tag{5.8}$$

One such widely used identification strategy is to set Q equal to the Choleski decomposition of Σ. Identification through a Cholesky decomposition implies making assumptions about the contemporaneous relationships between variables that are implicit in the ordering of the variables. In particular each variable responds contemporaneously only to variables that are ordered above it. These assumptions are often difficult to justify on theoretical grounds.

Blanchard and Quah (1989) proposed a method for distinguishing shocks based on whether they have a permanent or a temporary effect on the rate of growth of a nonstationary variable (explain). The long-run neutrality of money motivates one such restriction. In the context of our model, it states that shocks to the Federal Funds rate do not affect the steady state rate of growth of real GDP in the long run[37] so that the sum of the coefficients in the impulse response function should equal zero. These sums are given by the matrix $D(L)$ evaluated at $L = 1$:

$$D(1) = [I - B(1)]^{-1}Q \tag{5.9}$$

[35]They are all characterized by the same second moments properties.

[36]In addition to diagonalizing the variance-covariance matrix of the shocks, the variance of the shocks are normalized to unity.

[37]In the case where the measure of real activity is stationary (eg. CFNAI), then the long run restriction is stronger and implies that monetary policy does not affect the level of this variable in the long-run.

and the long-run restrictions are imposed by equating the relevant element in this “long-run” matrix to zero[38].

Although shocks can also be identified based on a structural equilibrium model where each equation in the system represents an equilibrium condition, VAR studies focusing on the impact of monetary shocks need only be semi-structural (Kilian, 2011) in that only the policy rate equation can be given a clear interpretation as the monetary policy rule. In particular, this argument has been used to justify ordering the policy variable after measures of real activity and inflation in the matrix of contemporaneous relationships[39]. In this case, only shocks in the policy rule equation need to be orthogonal to the other shocks and the matrix Q can be block-triangular instead of lower triangular (Christiano et al, 1999?; Den Haan et.al., 2007). Another popular identification scheme is through sign restrictions (Uhlig, 1997; Peersman and Wagner, 2014)[40].

Assuming we observe a significant response of one variable to a given shock, we may want to know how much of the overall behaviour of this variable can be explained by this kind of event. The forecast error variance decomposition provides such a measure. Using the VAR reduced-form estimates, we can recover the forecasts or conditional means of each variable given the current and past values of the variables included in the VAR and of the shocks. The forecast error variances or conditional variances measure the degree of uncertainty around these forecasts and depend only on the shocks that hit the system between the time the forecast is made and the time at which the variable is forecast:

$$\begin{aligned} \mathrm{var}_t(Y_{t+k}) &= \mathrm{var}_t(A(L)^{-1}\epsilon_t) = \sum_{j=0}^{k-1} A_j^{-1}\Sigma A_j^{-1'} \\ &= \mathrm{var}_t(A(L)^{-1}Q\upsilon_t) = \sum_{j=0}^{k-1} A_j^{-1}QQ'A_j^{-1'} \end{aligned} \tag{5.10}$$

(5.10) gives the k-step ahead forecast error covariance matrix. Any set of identifying assumptions (as embodied in Q) will also determine the percentage of the forecast error variance that can be attributed to each identified shock. The more of the forecast error variance of a variable can be attributed to a given structural shock, the more this shock can be considered to 'explain' the behaviour of the variable. Moreover, given the stationarity assumptions, the forecast and forecast error variance approach the unconditional mean and variance as the horizon of the forecast tends to infinity. Forecast error variance decompositions provide therefore interesting information at long horizons, when all the timing effects have played out, whereas impulse response functions are more informative about short and medium run dynamics.

[38] $B(1)$ is computed as the sum of the matrices of VAR coefficients at each lags: $\sum_{s=1}^{p}\Phi_s$ where p is the number of lags.

[39] An alternative view is that changes in the Feds Funds rate directly identifies policy innovations and should be ordered first (Bernanke and Blinder, 1992; Den Haan et. al., 2007).

[40] Other identification schemes exist (Romer and Romer, ?; Bernanke and Kuttner, 2005; Rigobon and Sacks, 2004; Bekaert et. al., 2013...).

Chapter 6

Bank portfolio choice and liquidity risk in the presence of nonlinearities

Abstract *In this chapter we empirically investigate the hypothesis that the composition of commercial banks' loan portfolio fluctuates according to conditions in financial markets. Commercial banks in the US have been increasingly concurrenced as main providers of both deposits and loans since the 1980s. Recent evidence suggests that commercial banks retain a competitive advantage in providing credit lines to businesses due to their ability to exploit the synergy between this activity and deposit-taking, both of them requiring banks to hold a buffer stock of liquid assets (Kashyap et.al., 2002). Moreover, because bank deposits can be seen as safe investments, liquidity shocks may give rise to simultaneous drawdowns of credit lines and a re-intermediation of funds via deposits, a flight to quality (Gatev and Strahan, 2006). Using weekly data on disaggregated US commercial banks' balance sheets over the period 1990-2007, we examine whether the CP-Tbill spread can help identify regimes where this hedge is effective in protecting bank loan supply from adverse financial market conditions using a 2-regime Threshold VAR approach. After finding significant evidence against linearity, we analyze the nonlinear propagation of shocks by computing Generalized Impulse Response Functions (GIRF). We find that commercial banks' loan portfolios are more likely to shift away from real estate loans and towards commercial and industrial loans when the paper-bill spread is low.*

Keywords: Liquidity risk, Flight to quality, Threshold VAR, GIRF

JEL Classification Numbers: C34, C52, G21, E44

6.1 Introduction

A number of studies point to the efficiency generated by the existence of a synergy between liquidity provision activities on both sides of the balance sheets of deposit-taking institutions. This synergy can in turn explain how resilient banks are to shocks to financial stability, both in terms of liquidity risk exposure and of their ability to sustain credit flows to the real economy.

Kashyap, Rajan and Stein (2002) discuss the conditions under which there exists a synergy between deposit-taking and the issuing of loan commitments or (unsecured) credit lines and to a broader extent term lending. Deposits are safe liquid short-term claims and loan commitments represent a form of liquidity insurance for borrowers, both exposing banks to liquidity risk. They argue that banks' unique ability to exploit this synergy may explain why banks have historically been offering these two services simultaneously. Both claims require the bank to hold a costly buffer stock of liquid assets. Assuming the liquidity needs of borrowers and depositors are imperfectly correlated, the overall cost of offering these two services can be reduced by engaging in these activities simultaneously. This provides a rationale for the observed tendency of deposit-taking banks to offer more (unsecured) credit lines. Bank heterogeneity renders this proposition difficult to test empirically. Simple OLS regressions across a sample of banks support the hypothesis that banks issuing more deposit-like claims hold a relatively larger stock of liquid assets and offer relatively more credit lines. These findings are consistent with the idea that any variation in the stock of liquid assets held by banks that is due to flows of deposit funding in or out of the banking system will be positively correlated with variations in loan commitments.

Gatev and Strahan (2006) focus on events such that the correlation between borrowers' and depositors' liquidity needs is negative (eg. 1998 events). This characterizes a situation where investors operate a flight to quality by pulling out of the commercial paper market and into safe bank deposits, thus forcing short-term corporate borrowers to draw on their credit lines at commercial banks. They further develop the asset pricing consequences of this theory. To the extent that banks' cost of funds is negatively related to the liquidity needs of borrowers (because they experience a simultaneous inflow of deposits) and that banks are competitive, they will be able to offer loan commitments at a lower price than any other non deposit-taking institution. Taking short-term changes in the CP-Tbill spread as a proxy for changes in the liquidity needs of firms (as reflected in the change in the cost of market-provided liquidity), they investigate how these changes affect the growth rates of banks' balance sheet items. Estimating weekly 3-variable VARs with the changes in the CP-Tbill spread, the Tbill rate and a balance sheet variable, they find that the former holds significant predictive power for both asset and liability side variables. On the asset side, both loans and liquid assets increase on impact following a widening of the spread, consistent with a drawdown of loan commitment associated with a simultaneous increase in funding. On the liability side, a rise in the spread leads to a surge in transaction deposits and a fall in banks' cost of funds (both actual and relative to finance companies' cost of funds). These effects do not appear to persist and subsequent periods are characterized by portfolio rebalancing movements.

Gatev, Schuermann and Strahan (2009) investigate further the implications of the model developed by Kashyap et. al. (2002) by examining directly the interaction between banks' deposit share, unused loan commitments and bank risk as proxied by the conditional volatility of bank equity returns. A higher deposit share signals a better hedge against liquidity risk stemming from the provision of liquidity to borrowers via loan commitments. Using a panel regression analysis with data for the 100 largest domestic banks between 1990 and 2002 and bank risk as the dependent variable, they identify how the mix between deposit share and loan commitment determines bank risk by interacting the two variables. They find that following a rise in loan commitments, the volatility of equity returns increase by more for banks with a lower share of transaction deposits. Moreover, this effect is stronger when the CP-TBill spread is larger. Note that the focus of the paper is on ex-post risk as opposed to ex-ante bank risk-taking as measured for example by the stock of liquid assets held by banks to protect themselves against liquidity shocks.

These portfolio choice decisions are explored in Cornett et. al. (2011) only for the period covering the financial crisis where tight liquidity conditions are proxied by the TED spread, measuring how difficult it is for banks to fund themselves in wholesale markets. In this paper liquidity risk also stems from exposure to wholesale deposit funding, the holding of illiquid assets such as asset-backed securities and leverage. They find that during the crisis, banks were more likely to increase their liquid asset holdings and reduce credit creation, following an increase in the TED spread, when they held a higher share of illiquid assets and a lower share of core deposits, capital and loan commitments on their balance sheets.

Acharya and Naqvi (2012) show how in some circumstances it can be a source of financial fragility. Abstracting from asset side liquidity risks, they build a model where banks allocate resources between loans and liquid assets which are used to weather shocks to deposit funding, the failure of doing so (liquidity shortfall) entailing a penalty cost. An agency problem between bank owners and bank managers yields an optimal compensation scheme whereby bank managers receive a bonus that is increasing in the volume of loans originated. This gives them an incentive to inflate the supply of loans (and asset prices). Because bank owners can credibly commit to auditing managers only in case of a liquidity shortfall and limited liability imposes an upper bound on the extent to which managers can be penalized, there will be a threshold amount of deposit funding above which liquidity risks are so small that managers will engage in excessive risk-taking by granting more loans. A rise in macroeconomic risk associated with rising credit spreads that generates a flight to quality (from bonds or commercial paper to bank deposits) may give rise to such a deterioration of agency problems within the banking sector.

This chapter differs in several respect from this literature. First, most of the literature is interested in variation in liquidity risk management strategies across banks. Here, we adopt a time series approach to investigate how commercial banks' ability to withstand liquidity shocks varies over time in the pre-crisis period. Second, we formally test for the existence of nonlinearities in the relationships between changes in a range of bank balance sheet variables and the paper-bill spread. Third, we use a threshold VAR approach to investigate whether bank protfolio changes in response to unexpected shocks to financial market conditions differ in regimes of high and low availability CP-Tbill spread for three different types of banks (large, small and foreign). Following Gatev and Strahan (2006), an increase in the former should be associated with an increase in liquidity availability (deposits funding) allowing banks to fund themselves more easily, and more so for high values of the CP-Tbill spread. We find significant evidence in favour of the hypothesis that the CP-Tbill spread can discriminate between periods of re-intermediation of funds allowing the traditional banking system to maintain the level of credit flows to the real economy and periods of disintermediation corresponding to a low-interest rate environment leaving commercial banks balance sheets more exposed to liquidity risks, which leads them to contract credit supply sharply.

The remaining of the chapter is organized as follows. In the next section, we discuss the relevance of threshold models as a way of modelling nonlinearities and present the data. In section 6.3, we describe the TVAR model, linearity tests and Generalized impulse response functions (GIRF). Section 6.4 presents the results using the CP-Tbill spread and the TED spread as threshold variables. Section 6.5 concludes.

6.2 Empirical strategy

6.2.1 Threshold nonlinearities

Threshold VAR models introduced by Tsay (1998) represent a parsimonious way of taking into account nonlinear dynamics. They assume that the relationships between the variables change "structure" when some key variable passes a threshold while allowing regime-switching to be endogenously determined. Hansen (1996) showed how to compute the empirical distribution for linearity tests, thus allowing one to conduct proper inference with respect to the presence of threshold nonlinearity. Koop et. al. (1996) showed how the presence of the threshold effect complicates the analysis of the propagation of shocks in such models. Because shocks to any variable may induce the system to change regime, the shape and magnitude of the impulse responses will depend on the sign and magnitude of the shock as well as the initial conditions in each regime, potentially implying the presence of asymmetric effects of shocks. Moreover, ignoring the possibility that future shocks will lead the system to switch to another regime (by setting all future realizations to zero as in the linear model) may cause impulse responses constructed in this way to understate the role of nonlinearities. The authors developed Generalized impulse response functions (GIRFs) as a solution to this problem. GIRFs are constructed as changes in forecast functions, conditioning the effect of each shock on the expected future history of shocks and on all possible initial conditions in each regime using simulation methods.

This methodology has been employed successfully to investigate the interactions between financial conditions and macroeconomic variables. For example, Balke (2000) and Atanasova (2003) estimate two-regime TVAR models for the US and the UK respectively, using credit spreads as threshold variable. Both analyses are theoretically grounded in models allowing the economy to switch to a credit rationing regime where financial conditions can have real effects and monetary policy becomes more efficient[1]. The empirical counterpart to this regime is one where credit spreads such as the CP-Tbill spread (proxying for the external finance premium) are high. Both find significant evidence in favour of threshold models against the linear alternative. In line with the theory, they find that monetary shocks have a stronger effect in the "credit constrained" regime. Balke (2000) also finds that contractionary monetary policy shocks have a nonnegligible marginal impact on the probability of switching from the "normal" to the "tight" credit regime.

In this chapter we are interested in the ability of commercial banks to attenuate adverse shocks to credit supply in these periods of financial turbulences. Kashyap et. al. (2002) provide a rationale for the coexistence of the two activities of liquidity provision to both borrowers and depositors within the same institution. Gatev and Strahan (2006) show that in times of strain in the commercial paper market, commercial banks are able to hedge their exposure to liquidity risk on the asset side of their balance sheet through their deposit-taking activity. Acharya and Naqvi (2011) develop a model where banks' portfolio decisions depend crucially on the overall funding liquidity available to them. When they benefit from easy credit conditions on the liability side they are more likely to grant loans.

This leads us to formulate the following hypothesis. In the high CP-Tbill spread regime, commercial banks benefit from easier access to funding liquidity (re-intermediation of funds) and this allows them to maintain or even expand the size of their balance sheets following adverse shocks to financial market conditions.

Our main credit spread measure is the CP-Tbill spread. Gatev and Strahan (2006) showed how this measure is more adapted to capture liquidity risk as opposed to credit risk, usually proxied by interest rate spreads between risky and safe bonds, such as the Baa-Aaa spread. Following them we investigate the response of balance sheet items to shocks to the CP-Tbill spread in both regimes to identify whether and how banks' liquidity risk exposure and portfolio decisions differ across regimes.

6.2.2 Data description

Weekly data on US commercial bank balance sheet is taken from the Federal Reserve Board's H.8 Release. From this database we obtain disaggregated data on both sides of the balance sheet for 3 categories of banks: large banks, small banks and foreign banks[2]. The sample period runs from January 1990 to August 2007. This covers a period characterized by a mature commercial paper market which also saw several episodes of financial stress such as the 1990-1991 Savings and Loans banking crisis and the 1998 financial turmoil that followed the failure of LTCM. Our sample period ends before the 2007-2009 financial crisis so as not to allow these events to bias our results. First, he crisis prompted the direct intervention of the Federal Reserve in some market for debt instruments, such as the commercial paper market starting in late 2008.

[1] See Blinder (1987) and Aziaradis and Smith (1998) for example.

[2] The data is estimated given information obtained from a sample of about 875 banks operating in the US. Out of these, large banks are the 25 largest by asset size, there are 60 reporting foreign banks and small banks make up the remaing ones (see http://www.federalreserve.gov/releases/H8/about.htm).

Second, a number of liquid assets such as mortgage-backed and asset-backed securities became illiquid in the early stages of the crisis, making it difficult to construct a consistent measure for liquid assets over the full period.

We analyze in turn the interaction between the paper-bill spread and several bank balance sheet variables. On the asset side, we are interested in loans and liquid assets. Loans are illiquid assets and we examine the two main subcategories of loans, commercial and industrial loans and real estate loans. The former have a shorter maturity and include loans drawn from credit lines. The latter are of a longer maturity and are to some extent more liquid as they may be securitized. We construct the liquid asset variable as the sum of securities (which contains treasury and agency securities, MBS and non-MBS securities), Fed Funds and reverse repos and cash assets. On the liability side, we look at demand deposits, large time deposits, borrowings from banks and nonbanks (including repo agreements). Following Gatev and Strahan (2006), we express balance sheet aggregates as the weekly changes normalized by the previous period total asset volume. In all models the 3-month Tbill rate represents the opportunity cost of funds. Our measure of liquidity need is the CP-TBill spread, the spread between the 3-month AA nonfinancial commercial paper rate and the 3-month Tbill rate. The Ted spread represents bank funding conditions in wholesale markets.

The Federal Reserve dataset presents several limitations. First it allows us to distinguish broadly between liquid and illiquid assets but not to identify liquidity risk exposure due to the issuance of credit lines. We do not have information about unused loan commitments or the total amount committed by banks which measure banks' exposure to liquidity risk on the asset side of their balance sheet. Cornett et. al. (2011) report that over the period 2006-2009, large banks were more exposed to this type of risk than small banks. Second, it does not allow to distinguish clearly the different type of funding sources on the liability side. We differentiate between core deposits, wholesale deposits (large time deposits) and borrowings (from banks and nonbanks). We do not have information on equity capital.

Table 6.1 describes how the structure of activity differs for the three categories of banks. On average, liquid assets and loans make up 30% and 60% respectively of the assets held by large and small banks on their balance sheet over the sample period. Commercial and industrial loans and real estate loans represent 15% and 30% of their asset portfolios. On average, slightly less than 60% of assets was funded by deposits and 15% by other debt instruments (borrowings). The absence of data on large time deposits for small banks over most of the sample suggest that, contrary to large banks, they have less recourse to this source of funding. Foreign banks differ from domestic banks on several respects. First, portfolio shares appear to be more volatile. Second, they hold a larger share of their assets in liquid form and a lower share of loans (39% and 55% respectively), the latter being comprised mainly of commercial and industrial loans (30%). Most of these are funded by large time deposits (44%) and other debt instruments (32%). Thus foreign banks are more likely to be exposed to liquidity shocks on either side of the balance sheet. To protect themselves against this heightened exposure to liquidity risk, they hold a larger buffer stock of liquid assets. Since foreign banks have access to outside resources in the face of liquidity risk, including them in the analysis may not be very informative.

Table 6.2 reports summary statistics on the variables used in the empirical analysis. On average all balance sheet items grew over the sample period. Changes in liquid asset holdings are more volatile, as are changes in deposits for domestic banks.

6.3 Methodology

6.3.1 Specification of the TVAR model

We specify the 2-regime TVAR model in recursive form as follows:

$$Y_t = C_1 + A_1 Y_t + B_1(L) Y_{t-1} + (C_2 + A_2 Y_t + B_2(L) Y_{t-1}) I(q_{t-d} > \gamma) + \epsilon_t \qquad (6.1)$$

where Y_t is the vector of endogenous variables $[\Delta BS \quad \Delta TB \quad CPTB]'$ where ΔBS is the weekly change in the balance sheet variable, ΔTB is the weekly change in the 1-month TB rate and $CPTB$ is the spread between the 1-month CP rate and the 1-month TB rate. C_1 and C_2 are vectors of constant parameters, $B_1(L)$ and $B_2(L)$ are matrices of lag polynomials of order p, $I(.)$ is an indicator function that takes on a value of 1 when the threshold variable q_{t-d} is below the threshold value γ and 0 otherwise, ϵ_t is a vector of residuals with covariance matrix Σ_ϵ, d is the delay parameter determining the timing of the switch in regime. The threshold variable q_{t-d} is the lagged value of the CP-TBill spread, which is an endogenous variable in the system.

Specification (6.1) allows the contemporaneous relationships represented by the matrices of coefficients A_1 and A_2 to change structure in each regime in addition to the lag polynomial coefficients and the constant terms. Because of its greater tractability (and following the previous literature), we impose a recursive structure on these matrices:

$$A_i = \begin{bmatrix} 0 & 0 & 0 \\ \bullet & 0 & 0 \\ \bullet & \bullet & 0 \end{bmatrix}$$

This structure identifies the estimated shocks as orthogonal ("structural") innovations although it is difficult to justify this choice at a weekly frequency. As slow-moving variables such as bank credit can be assumed to react with a lag to financial variables, we order changes in bank balance sheet data first followed by the change in the TB rate and the CP-TB spread.

Specification (6.1) does not allow the variance-covariance matrix of the structural residuals to differ across regimes (regime-dependent heteroskedasticity)[3]. The threshold VAR can be described as a linear VAR with time-varying parameters, alternating between two regimes according to a simple switching rule that is triggered by the behaviour of the threshold variable. For greater tractability, the number of lags is not allowed to vary across regimes.

[3]Note that other forms of heteroskedasticity might be present in the data both in the linear and threshold model, which we can take into account when testing for a threshold effect.

6.3.2 Estimation and specification tests

To estimate model (6.1) we adopt the following strategy[4]. First, for given lag length p, the model parameters are estimated by OLS. The estimated values of (γ, d) minimizes the log determinant of the structural residuals:

$$(\widehat{\gamma}, \widehat{d}) = \underset{(\gamma,d)\, \in\, (S\times D)}{\arg\min} \quad \log|\widehat{\Sigma}(\gamma, d)| \tag{6.2}$$

Γ is the set of all threshold values used in the grid search[5] and $D = [1, p]$. Second, we repeat the procedure for all $p \in P = [1, 10]$ and examine those favoured by the AIC and SIC criteria[6].

Next, having estimated the threshold TVAR model, we want to know whether it performs statistically better than the linear model. A conventional method to test the statistical significance of the threshold nonlinearity is to construct a Wald test of the joint restrictions $H_0 : C_2 = 0, A_2 = 0, B_2(L) = 0$ in (6.1)[7]. If the parameter γ was known a priori, the Wald test would follow a χ^2 distribution. But since we estimate this parameter alongside the others, traditional inference methods are not valid because γ is not defined under the null hypothesis[8]. Davies (1977) developed a test for which inference can be conducted without prior knowledge of the two parameters, the Sup Wald statistics:

$$\text{Sup}W = \max_{\gamma} W(\gamma) \tag{6.3}$$

To make inference, we use the algorithm proposed by Hansen (1996) to simulate asymptotic p-values for such tests[9]. These results rely on the assumption of stationarity and ergodicity of the TVAR process. A sufficient condition for stability is that the maximum eigenvalue of the characteristic polynomial of the linear VAR process in both regimes be inferior to one in absolute value. Finally, if the errors in the TVAR cannot be assumed to be conditionally homoskedastic, the test statistics allows for a general form of heteroskedasticity (see Appendix A).

Equation-by-equation tests of nonlinearity may yield additional information about the nonlinear dynamics of the system by indicating whether some variables are more sensitive to the threshold effect and whether the threshold estimate is consistent across equations.

[4] A similar procedure was employed by Atanasova (2003), Calza and Sousa (2006) and Mandel (2012).

[5] Following Hansen (1996), Γ corresponds to the range of values taken by the threshold variable trimmed at the bottom and the top by 15% plus the maximum number of coefficients estimated in the system of equations.

[6] The AIC and SIC criteria are computed as follows: AIC $= \ln|\widehat{\Sigma}(\gamma)| + \frac{2M}{T}$ and SIC $= \ln|\widehat{\Sigma}(\gamma)| + \frac{\ln TM}{T}$ where M is the number of unrestricted parameters in the TVAR.

[7] See Appendix A. We run these test for the chosen lag length and delay parameter.

[8] As an example, assume homoskedasticity of the error terms, then the Wald test $W(\gamma)$ reaches its maximum when $\gamma = \hat{\gamma}$ by construction. Therefore our test statistics $W(\hat{\gamma})$ is the random maximum of the random function $W(\gamma)$ whose arguments each follow a χ^2 distribution and its distribution lies to the right of the χ^2 distribution, thus increasing the probability of making a Type-I error.

[9] Hansen (1999) discusses the comparative advantages of simulating their bootstrap distribution and notes that their is no clear evidence as to which method provides the best approximation although he recommends using the latter. See Appendix B

6.3.3 Generalized impulse response functions (GIRFs)

Contrary to its linear counterpart, nonlinear models imply that the response of variables to shocks are history and shock dependent. Because shocks may induce the system to switch regime, we need to assess the effect of the shock on the probability of regime-switching in addition to the magnitude, direction and persistence of the responses. Koop, Pesaran and Potter (1996) developed one approach to deal with these problems, consisting in averaging out the effects of future shocks. The idea behind the Generalized Impulse Response function (GIRF) can be illustrated in the case of the threshold VAR model. Assuming the linear and threshold models take the simple reduced-forms:

$$Y_t = BY_{t-1} + \delta_t$$
$$Y_t = B_1 Y_{t-1} + B_2 Y_{t-1} I(q_{t-d} > \gamma) + \delta_t$$

Let $\{\delta_t^0\}_{t=0}^n$ be a sequence of shocks taken randomly from the joint distribution of the VAR residuals. Let $\{\delta_t^*\}_{t=0}^n$ be the same sequence except for δ_0^* which differs from the former only in the value of its i^{th} component that is exogenously fixed to be equal to η. Then to study the effect of a shock η on the time profile of the time series, generalized impulse responses at horizon n (conditioning on a particular realization of Y_{t-1}) can be constructed as follows in the linear case:

$$GIR_Y(n, \eta, Y_{t-1}) = E[Y_{t+n}^*] - E[Y_{t+n}^0] = B^{n+1}\delta_0^*$$

which is independent of the initial condition Y_{t-1} and of the future sequence of shocks $\{\delta_t^0\}_{t=1}^n$ as well as from the size and magnitude of η (up to a scale parameter). In the threshold case, the generalized impulse responses become:

$$\begin{aligned} GIR_Y(n, \eta, Y_{t-1}) = E\left\{\prod_{i=0}^{n}\left(B_1 + B_2 I(q_{t+i-d}^* > \gamma)\right) - \prod_{i=0}^{n}\left(B_1 + B_2 I(q_{t+i-d}^0 > \gamma)\right)\right\} Y_{t-1} \\ + E\left[\sum_{j=0}^{n}\left\{\prod_{i=j+1}^{n}\left(B_1 + B_2 I(q_{t+i-d}^* > \gamma)\right)\right\}\delta_j^*\right] - E\left[\sum_{j=0}^{n}\left\{\prod_{i=j+1}^{n}\left(B_1 + B_2 I(q_{t+i-d}^0 > \gamma)\right)\right\}\delta_j^0\right] \end{aligned}$$

which depends on the initial conditions Y_{t-1} and the *expected* future paths of the threshold variable $\{q_t^*\}_{t=1}^n$ and $\{q_t^0\}_{t=1}^n$. This difference rests entirely on the possibility of regime-switching and the endogeneity of the threshold variable. Note that the distribution of the threshold variable q_{t-d}^* depends itself on the initial conditions Y_{t-1}, the entire sequence of future shocks $\{\delta_t^0\}_{t=0}^n$ as well as the sign and magnitude of the initial shocks η. This feature generates two sources of asymmetries. First, depending on the initial regime the economy is in (ie the initial value of q_{t-d}), a strong threshold effect will be associated with different dynamics of the system between the two regimes[10]. Because initial conditions in each regime matter, GIRFs are usually computed conditioning on the initial regime the economy is in and then averaging over

[10] Note however that impulse responses in the two regimes are not independently distributed because of the possibility of regime switching

all the possible histories within each regime[11]. Second, shocks of different sign and magnitude may be expected to have asymmetric effects within each regime. Given these features of the threshold model, a key difficulty is finding the adequate theoretical counterpart to these forms of asymmetries. The former relates to structural shifts in the relationships between key variables, whereas the latter to threshold effects in the presence of regime-switching. Even if investigating endogenous regime-switching is not a primary motive of the analysis, controlling for it may provide a more realistic estimate of the impulse responses to shocks when one investigates threshold effects. For these reasons, we control for future regime switches by bootstrapping the residual series from the estimated TVAR and use it to simulate the path of the system (benchmark series) instead of setting all future shocks to zero. To generate the impulse response to a shock to variable i, we simulate the path of the system when replacing the shock to variable i in period 0 by a η (1,-1, 2,-2) standard deviation shock and substract the benchmark series. To construct the GIRF we repeat this procedure 500 times and take the average over the bootstrap simulations (See appendix C for details).

$$GIR_Y(n, \eta, R) = E[Y^*_{t+n}|\eta, R] - E[Y^0_{t+n}|R] \tag{6.4}$$

where R indicates the initial regime, $\{Y^*\}_{t=1}^n$ is the simulated series imposing shock η in period 0 and $\{Y^0\}_{t=1}^n$ is the benchmark series.

One important limitation of this approach is that impulse responses are generated assuming the true model is known, in other words it ignores uncertainty about the choice of the nonlinear model, as well as sampling variability with respect to the model's parameter estimates. In addition, it is computationally difficult to accomodate the presence of heteroskedasticity in the error terms using this methodology.

6.4 Results

6.4.1 Parameter estimation and linearity tests

We proceed with the analysis by examining models with the CP-Tbill spread examining in turn its relationship with each balance sheet variable for each of the three types of banks. First we run a set of tests in order to select the lag length (p) and the delay parameter (d) used in the TVAR specification. Tables 6.3-6.5 present the estimated parameters and residual tests for the specifications chosen by the AIC and SIC criteria and for the models with large banks, small banks and foreign banks respectively. Columns 5 to 10 of each table give the p-values for equation-by-equation residual tests of serial correlation and conditional heteroskedasticity (Tsay, 1998; Hubrich and Terasvirta, 2013)[12].

We observe that the model chosen by the SIC criterion, giving an optimal lag length of 1 and delay parameter of 2 weeks in most specifications, outperforms the one chosen by the AIC

[11] By averaging over all possible initial conditions in each regime we give equal weights to each of them.

[12] Residual tests are standard LM tests constructed by regressing each equation residuals and squared residuals respectively on 12 lagged values of themselves (See Lutkepohl (2006)).

criterion in that it shows almost no residual serial correlation (columns 6 to 8). The table also shows that the residuals from the balance sheet equations exhibit significant ARCH effects[13]. Since the model with one lag is also the most parsimonious model, we choose this specification to estimate the threshold and test for linearity. Results for these are presented in Tables 6.7, 6.9 and 6.10. Panel A presents the test statistics constructed using the coefficients in all three equations (full VAR). The null hypothesis of linearity is rejected at the 10% significance level in most specifications. Panels B to D present the test-statistics for equation-by-equation tests of linearity with the CP-Tbill spread as threshold variable. While we are usually unable to reject the null of linearity for the balance sheet equations (except for foreign banks), it is largely rejected for those for the change in the Treasury Bill rate and the CP-Tbill spread. Although the threshold estimate varies from one specification to another (first rwo in each panel), the value of 0.26 is the most frequent and is consistent across the three groups of banks. This is the value we will retain in the remaining of the analysis. Figure 6-2(a) depicts the CP-Tbill spread along with periods of low and high spread regimes. We can distinguish two main periods in which the economy was in the low spread regime. The first one covers most of 1992-1993 and the second spans the period between mid-2001 and mid-2005. Both periods were characterized by a low-interest rate environment (low Feds Funds rate)) and reduced volatility of the paper-bill spread.

6.4.2 Impulse response analysis

Figure 6-2 presents the impulse responses estimated using a linear specification. A positive CP-Tbill spread shock leads to temporary increases in loans (including real estate loans), large time deposits and borrowings and a persistent increase in commercial and industrial loans. The responses of liquid assets and deposits are not significant. The finding that bank loan supply increases following a positive shock to the paper-bill spread is consistent with the conjecture that a liquidity shock in the commercial paper market leads to a reintermediation of funds, allowing commercial banks to expand their lending. Inspection of the GIRFs reveals that the assumption of linearity may hide significant differences across regimes.

Figures 6-3-6-5 presents the Generalized impulse responses to an 8 basis point shock to the CP-Tbill spread in both the high and low CP-Tbill spread regimes. In addition, linear impulse responses are represented on the same graph. The fact that this shock is relatively small compared to a standard deviation of 24 basis points and an interquartile range of 27 basis points for the spread over the sample reflects our focus on within regime variation rather than on the nonlinear propagation of shocks.

Figure 6-3 shows the impulse response of the balance sheet variables of large banks. Following a positive shock to the CP-Tbill spread, large banks experience a fall (rise) in total assets and liabilities in the low (high) spread regime and a rise in commercial and industrial loans in both regimes. This finding is consistent with the idea that a rise in the CP-Tbill spread leads firms to draw on their credit lines with banks, thus leading to a rise in commercial and

[13] How does this impact the future analysis?

industrial loans. It is also consistent with the hypothesis that in the high spread regime, funding liquidity availability is higher for banks because of a flight to quality effect, which enables them to expand their balance sheets, or at least not contract credit much in the face of adverse circumstances in financial markets (as in the low spread regime).

The low spread regime differentiates itself from the high spread regime in several respects. First, the magnitude of the impact on commercial and industrial loans is seven times larger in the low spread regime, rising up to 0.014%, to be compared with a mean percentage growth of 0.010% over the period, which is equivalent to a doubling of the growth rate of C&I loans. In the high spread regime the increase amounts to only 0.002%. Second, in the low spread regime the shock to the paper-bill spread occasions a fall in the growth rate of real estate loans from an average of 0.04% to a little over 0.03% whereas in the high spread regime it slightly increases on impact. Third, the positive shock to the paper-bill spread leads to a fall in deposits of 0.06% on impact, compared to an average growth rate of 0.04%, so that deposits fall in absolute terms. Thus, in the low spread regime, a rise in the spread also leads depositors to withdraw money from banks, the opposite of a flight to quality. This lead them to contract their loan portfolios (and more generally credit) and to draw on their stock of liquid assets. In contrast, in the high spread regime the positive shock to the paper-bill spread leads to a slight increase in deposit growth. Fourth, the growth rate of liquid assets becomes negative, with a fall of 0.06% on impact to be compared to an average growth rate of 0.03%. This is consistent with banks drawing on their stocks of liquid assets when liquidity shortages arise on both sides of their balance sheets. In the high spread regime, the effect of a spread shock on liquid assets is positive. Finally, the impulse responses of C&I loans, real estate loans, liquid assets and deposits are more persistent in the low spread regime, so the impact on the composition of loan portfolios is more durable. In particular, although we observe portfolio rebalancing effects in the second period with sharp increases in liquid assets and deposits, these are not enough to make up for the outflow of deposits and of liquid assets that occurred on impact.

Impulse responses for small and foreign banks (Figures 6-4 and 6-5) are roughly similar. One notable difference is that these banks are partly compensating the loss of deposit funding through either higher borrowings (small banks) or inflows of large time deposits (foreign banks) in the low spread regime. However in both cases we still observe large outflows of liquid assets. Foreign banks also differ in that because of the higher share C&I loans on their balance sheets, total credit (loans) rises in the low spread regime following a positive shock to the paper-bill spread although we still observe the compositional shift towards more business loans and less real estate loans.

Overall these findings suggest that banks contract their balance sheets following a positive shock to the CP-Tbill spread in the low spread regime as predicted by theories of financial frictions, but not in the high spread regime. These findings suggest that commercial banks can alternatively smooth shocks to credit supply when the paper-bill spread is high and amplify lending cycles when it is low. An important factor that may interfere with our results is the possibility that credit demand also varies across regimes.

6.4.3 An alternative measure of financial conditions

We also conduct the analysis using the TED spread, which is a widely used indicator of bank funding conditions in wholesale markets using data for large banks only. An increase in the TED spread reflects an increase in the cost of funds in wholesale markets and may also incite borrowers to draw on their credit lines if they perceive that banks' ability to extend credit may be reduced in the near future. Since the TED spread signals increased counterparty risk, it cannot be associated with a flight to quality contrary to the paper-bill spread. Linearity tests reported in Table 6.8 shows that evidence in favour of threshold nonlinearity is less strong than with the paper-bill spread. The most likely threshold estimate is 0.66. The TED spread is depicted in figure 6-2(b) with the shaded area representing the high spread regime. This regime mainly characterises spikes in the TED spread, corresponding to episodes of extreme strain in the financial sector, which renders comparison with previous results difficult. Nevertheless, Generalized impulse responses to an 8 basis point shock to the TED spread in both regimes are represented in Figure 6-6. They exhibit responses that are smaller in magnitude and with no clear difference across regimes for loans and deposits. This results comfort the hypothesis that changes in the CP-Tbill spread can uniquely identify episodes of re-intermediation of funds during period of financial turmoil.

6.5 Conclusion

Employing a threshold VAR methodology, we investigate the presence and the role of nonlinear relationships between changes in the CP-Tbill spread, the Treasury Bill rate and commercial banks' balance sheet variables in the US in the run-up the the 2007-2009 financial crisis. The former has been extensively used in the literature on the financial accelerator as a gauge of the extent of frictions in financial markets. In the banking literature, increases in the paper-bill spread have been shown to signal the simultaneous drawdown of credit lines by borrowers and the re-intermediation of funds into bank deposits by risk-averse investors. In this chapter, we make several contributions to the literature. We are the first to formally test and find significant evidence in favour of threshold linearity with respect to the level of the paper-bill spread in the relationship between the latter and bank balance sheet variables. Moreover, the estimated threshold is consistent across a range of models including both asset and liability side variables and across categories of banks (large, small and foreign). It clearly differentiates a low spread regime characterized by low volatility and broadly corresponding to periods of low policy rates.

Second, computing Generalized impulse responses to shocks to the paper-bill spread, we show that commercial banks are more likely to reduce credit supply in the low spread regime where we observe sharp increases in commercial and industrial loans acompanied by sharp withradawals of deposits and falls in real estate loans and liquid assets. On the contrary, in the high spread regime, deposits and liquid assets remain stable or increase, allowing banks to hedge the liquidity risk exposure due to the unexpected drawdown of credit lines. This result extends previous findings that banks with a higher share of deposits reduced less credit supply during the financial crisis than the others (Cornett. et. al.). Our findings suggest that periods of very low paper-bill spread are associated with a disintermediation of funds that may have significant effects on the ability of commercial banks to respond to financial disruptions. The negative correlation between deposit withdrawals and credit line drawdowns may be less pronounced in those periods, where the safe asset property of bank deposits is less valuable. As a result it is harder for banks to attract deposit funding when a liquidity risk on the asset side of their balance sheets materializes.

Table 6.1: Summary statistics: bank characteristics

The table reports summary statistics for bank characteristics as reflected by the percentage share of liquid assets, loans, commercial & industrial loans and real estate loans on the asset side and deposits, large time deposits and bank borrowings on the liability side for large, small and foreign banks. Balance sheet data for a panel of commercial banks operating in the US is observed weekly and seasonally adjusted over the period 4-Jan-1989 to 1-Aug-2007. The data is from the Federal Reserve Bank H.8 Release. On the asset side, TA stands for total assets, $LOAN$ are total loans, CIL are commercial and industrial loans, REL are real estate loans and LIQ are liquid assets (sum of securities, Fed Funds and reverse repos and cash assets). On the liability side, DEP are demand deposits, LTD are large time deposits, BOR is credit obtained from banks and others (incl. Fed Funds and repurchase agreements). The table displays the first (Q1) and third (Q3) quantiles as well as the median, the mean and the standard deviation of the distribution of the variables over the sample period.

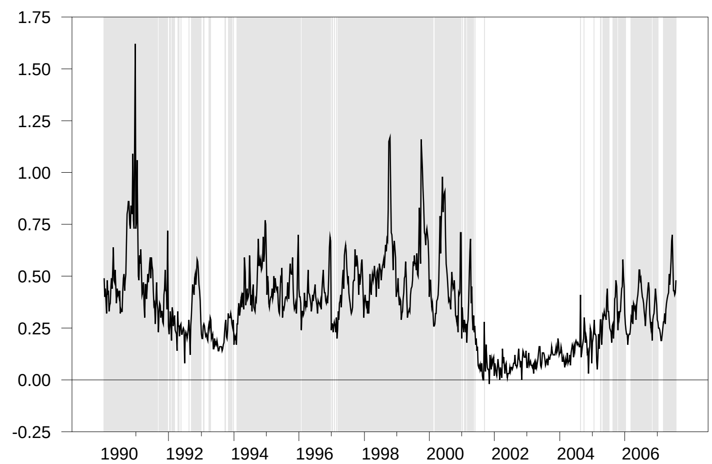
	Mean	Q1	Median	Q3	Sd
Panel A: Large banks characteristics					
$LIQ_t/Assets_t$	30.84	29.24	30.29	31.99	2.46
$LOAN_t/Assets_t$	60.95	59.46	61.17	62.33	1.62
$CIL_t/Assets_t$	14.12	12.42	14.24	15.50	2.02
$REL_t/Assets_t$	27.39	26.24	26.75	27.67	2.18
$DEP_t/Assets_t$	58.96	52.30	58.80	65.18	7.12
$LTD_t/Assets_t$	8.30	7.42	8.10	8,96	1.58
$BOR_t/Assets_t$	15.38	12.43	16.04	17.64	2.92
Panel B: Small banks characteristics					
$LIQ_t/Assets_t$	31.58	28.39	30.14	34.92	4.54
$LOAN_t/Assets_t$	65.97	64.62	66.26	67.88	2.50
$CIL_t/Assets_t$	15.56	13.26	15.93	17.01	2.97
$REL_t/Assets_t$	35.75	30.78	33.71	40.61	5.95
$DEP_t/Assets_t$	57.55	55.12	57.34	59.59	3.19
$BOR_t/Assets_t$	14.82	13.30	14.67	16.61	2.06
Panel C: Foreign banks characteristics					
$LIQ_t/Assets_t$	38.72	31.32	39.26	46.86	10.91
$LOAN_t/Assets_t$	54.84	51.95	54.74	57.33	4.26
$CIL_t/Assets_t$	29.96	24.12	32.21	35.99	7.97
$REL_t/Assets_t$	6.02	2.51	3.63	9.63	4.20
$DEP_t/Assets_t$	3.12	2.25	2.78	4.04	1.00
$LTD_t/Assets_t$	43.91	30.81	40.93	58.02	15.26
$BOR_t/Assets_t$	31.75	23.83	29.00	40.99	9.25

Table 6.2: Summary statistics

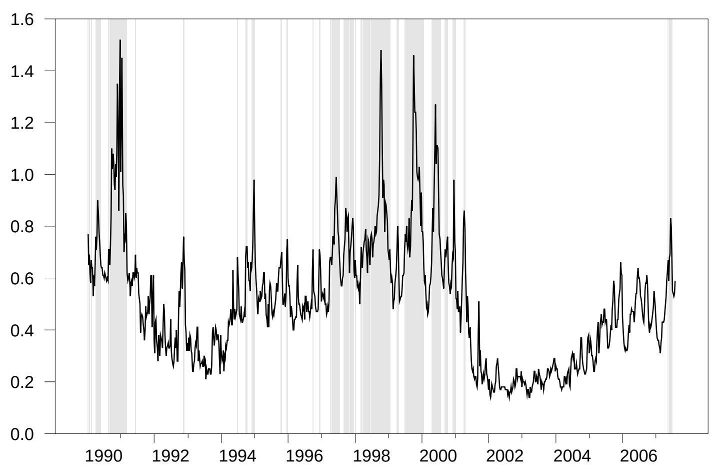
The table reports summary statistics for changes in liquid assets, total loans, commercial & industrial loans, real estate loans, deposits, large time deposits and bank borrowings for large, small and foreign banks. Changes are expressed in percentage points and normalized by the previous period total assets. Balance sheet data for a panel of commercial banks operating in the US is observed weekly and seasonally adjusted over the period 4-Jan-1989 to 1-Aug-2007. The data is from the Federal Reserve Bank H.8 Release. On the asset side, TA stands for total assets, $LOAN$ are total loans, CIL are commercial and industrial loans, REL are real estate loans and LIQ are liquid assets (sum of securities, Fed Funds and reverse repos and cash assets). On the liability side, DEP are demand deposits, LTD are large time deposits, BOR is credit obtained from banks and others (incl. Fed Funds and repurchase agreements). TB represents the 3-month TBill rate (FRED), $CPTB$ is the spread between the 3-month AA non-financial commercial paper rate (Global Financial database) and the 3-month TBill rate. TED is the TED spread (FRED), the difference between the 3-month LIBOR rate and the 3-month TBill rate. BAA is the spread between the Baa bond rate and the Aaa bond rate (FRED). The table displays the first (Q1) and third (Q3) quantiles as well as the median, the mean and the standard deviation of the distribution of the variables over the sample period.

Label	Construction	Mean	Q1	Median	Q3	Sd
		Panel A: Large banks				
ΔLIQ	$\Delta LIQ_t/Assets_{t-1}$	0.03	-0.21	0.02	0.27	0.40
$\Delta LOAN$	$\Delta LOAN_t/Assets_{t-1}$	0.06	-0.04	0.07	0.18	0.20
ΔCIL	$\Delta CIL_t/Assets_{t-1}$	0.01	-0.02	0.01	0.04	0.06
ΔREL	$\Delta REL_t/Assets_{t-1}$	0.04	-0.01	0.04	0.09	0.13
ΔDEP	$\Delta DEP_t/Assets_{t-1}$	0.04	-0.19	0.04	0.27	0.40
ΔLTD	$\Delta LTD_t/Assets_{t-1}$	0.01	-0.03	0.01	0.04	0.06
ΔBOR	$\Delta BOR_t/Assets_{t-1}$	0.02	-0.12	0.03	0.16	0.25
		Panel B: Small banks				
ΔLIQ	$\Delta LIQ_t/Assets_{t-1}$	0.04	-0.31	0.04	0.39	0.71
$\Delta LOAN$	$\Delta LOAN_t/Assets_{t-1}$	0.11	-0.06	0.11	0.25	0.39
ΔCIL	$\Delta CIL_t/Assets_{t-1}$	0.02	-0.03	-0.01	0.06	0.12
ΔREL	$\Delta REL_t/Assets_{t-1}$	0.08	-0.00	0.07	0.14	0.22
ΔDEP	$\Delta DEP_t/Assets_{t-1}$	0.10	-0.26	0.08	0.42	0.67
ΔBOR	$\Delta BOR_t/Assets_{t-1}$	0.02	-0.25	0.02	0.31	0.59
		Panel C: Foreign banks				
ΔLIQ	$\Delta LIQ_t/Assets_{t-1}$	0.09	-0.55	0.10	0.75	1.06
$\Delta LOAN$	$\Delta LOAN_t/Assets_{t-1}$	0.09	-0.29	0.10	0.47	0.66
ΔCIL	$\Delta CIL_t/Assets_{t-1}$	0.03	-0.11	0.02	0.15	0.25
ΔREL	$\Delta REL_t/Assets_{t-1}$	0.00	-0.04	-0.00	0.04	0.10
ΔDEP	$\Delta DEP_t/Assets_{t-1}$	0.00	-0.04	0.00	0.05	0.09
ΔLTD	$\Delta LTD_t/Assets_{t-1}$	0.11	-0.22	0.11	0.47	0.63
ΔBOR	$\Delta BOR_t/Assets_{t-1}$	0.06	-0.66	0.07	0.70	1.19
		Panel D: Interest rates				
ΔTB	-	-0.00	-0.04	0.00	0.04	0.08
$CPTB$	-	0.38	0.22	0.36	0.49	0.24
TED	-	0.54	0.33	0.50	0.67	0.28
BAA	-	0.84	0.68	0.82	0.94	0.20

Figure 6-1: Plot of the credit spreads



(a) CP-Tbill spread



(b) TED spread

Note: The CP-Tbill spread is the difference between the 3-month AA nonfinancial commercial paper rate and the 3-month Tbill rate (Global Financial data). The TED spread is the difference the 3-month LIBOR rate and the 3-month Tbill rate (FRED). The shaded areas represent the high spread regimes estimated in the empirical analysis (the estimated thresholds are 0.26 and 0.66 respectively).

Table 6.3: Parameter selection: Large banks/CP-Tbill spread

This table presents the parameters (p and d) and residual tests associated with the specifications of the TVAR chosen by the AIC and SIC criteria respectively. We consider specifications with up to 4 lags and delay up to 4 weeks. The models estimated are 3-variable TVAR models including the changes in bank balance sheet variables (ΔBS), the Tbill rate (ΔTB) and the commercial paper-Tbill spread ($CPTB$) in that order. The threshold variable is the spread variable $CPTB$. The data is weekly from 1-Jan-1990 to 1-Aug-2007. ΔTB is the first difference of the annualized 3-month Tbill rate (FRED). $CPTB$ is constructed as the 3-month AA (non-financial) commercial paper rate minus the 3-month Tbill rate (Global Financial data). Commercial bank balance sheet data is from the Federal Reserve Bank H.8 Release. All series are seasonally adjusted. On the asset side, TA stands for total assets, $LOAN$ are total loans, CIL are commercial and industrial loans, REL are real estate loans and LIQ are liquid assets (sum of securities, Fed Funds, reverse repos and cash assets). On the liability side, DEP are demand deposits, LTD are large time deposits, BOR is credit obtained from banks and others (incl. Fed Funds and repurchase agreements). Changes are normalized and expressed in percentage terms as described in Table 6.2.

				Autocorrelation tests			ARCH tests		
	TVAR	p	d	ΔBS	ΔTB	$CPTB$	ΔBS	ΔTB	$CPTB$
				$\Delta BS = \Delta LIQ$					
AIC	-12.73	4	3	**0.000**	0.537	**0.000**	0.724	**0.000**	**0.000**
SIC	-12.28	1	3	**0.000**	**0.000**	**0.000**	0.327	**0.000**	**0.000**
				$\Delta BS = \Delta LOAN$					
AIC	-13.44	4	3	0.200	0.577	**0.000**	1.000	**0.000**	**0.000**
SIC	-13.03	1	3	**0.056**	**0.000**	**0.000**	0.997	**0.000**	**0.000**
				$\Delta BS = \Delta CIL$					
AIC	-16.24	4	2	**0.032**	0.574	**0.004**	0.971	**0.000**	**0.000**
SIC	-15.78	1	2	**0.000**	**0.000**	**0.000**	0.578	**0.000**	**0.000**
				$\Delta BS = \Delta REL$					
AIC	-14.56	4	2	0.864	0.628	**0.000**	1.000	**0.000**	**0.000**
SIC	-14.15	1	2	0.136	**0.000**	**0.000**	1.000	**0.000**	**0.000**
				$\Delta BS = \Delta DEP$					
AIC	-12.79	4	4	**0.000**	0.654	**0.000**	0.936	**0.000**	**0.000**
SIC	-12.32	2	4	**0.000**	**0.002**	**0.000**	0.991	**0.000**	**0.000**
				$\Delta BS = \Delta LTD$					
AIC	-16.39	4	2	**0.000**	0.489	**0.000**	**0.000**	**0.000**	**0.000**
SIC	-15.92	1	2	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
				$\Delta BS = \Delta BOR$					
AIC	-13.40	4	2	0.157	0.549	**0.000**	**0.011**	**0.000**	**0.000**
SIC	-12.96	1	2	**0.009**	**0.000**	**0.000**	**0.046**	**0.000**	**0.000**

Note: The table displays the minimum value of the criteria achieved over values of $p \in [1, 4]$ and $d \in [1, 4]$ and displays the associated values of p and d for the trheshold model. Residual autocorrelation and ARCH tests are LM tests based on 12 lags applied to the specifications described in columns 3 to 5. Only p-values for the tests on the residuals of each single equation in the VAR are reported. In bold are the test-stats rejecting the null hypothesis (at the 10% significance level) of serial correlation and heteroskedasticity.

Table 6.4: Parameter selection: Small banks/CP-Tbill spread

This table presents the parameters (p and d) and residual tests associated with the specifications of the TVAR chosen by the AIC and SIC criteria respectively. We consider specifications with up to 8 lags and delay up to 8 weeks. The models estimated are 3-variable TVAR models including the changes in bank balance sheet variables (ΔBS), the Tbill rate (ΔTB) and the commercial paper-Tbill spread ($CPTB$) in that order. The threshold variable is the spread variable $CPTB$. The data is weekly from 1-Jan-1990 to 1-Aug-2007. ΔTB is the first difference of the annualized 3-month Tbill rate (FRED). $CPTB$ is constructed as the 3-month AA (non-financial) commercial paper rate minus the 3-month Tbill rate (Global Financial data). Commercial bank balance sheet data is from the Federal Reserve Bank H.8 Release. All series are seasonally adjusted. On the asset side, TA stands for total assets, $LOAN$ are total loans, CIL are commercial and industrial loans, REL are real estate loans and LIQ are liquid assets (sum of securities, Fed Funds, reverse repos and cash assets). On the liability side, DEP are demand deposits, LTD are large time deposits, BOR is credit obtained from banks and others (incl. Fed Funds and repurchase agreements). Changes are normalized and expressed in percentage terms as described in Table 6.2.

				Autocorrelation tests			ARCH tests		
	TVAR	p	d	ΔBS	ΔTB	$CPTB$	ΔBS	ΔTB	$CPTB$
				$\Delta BS = \Delta LIQ$					
AIC	-11.77	4	2	**0.000**	0.782	**0.001**	**0.000**	**0.000**	**0.000**
SIC	-11.34	1	3	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
				$\Delta BS = \Delta LOAN$					
AIC	-12.50	4	1	**0.015**	0.542	**0.000**	**0.071**	**0.000**	**0.000**
SIC	-12.05	1	2	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
				$\Delta BS = \Delta CIL$					
AIC	-14.87	4	2	**0.045**	0.596	**0.005**	**0.000**	**0.000**	**0.000**
SIC	-14.46	1	2	**0.008**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
				$\Delta BS = \Delta REL$					
AIC	-13.58	4	1	0.103	0.732	**0.000**	1.000	**0.000**	**0.000**
SIC	-13.16	1	2	**0.004**	**0.000**	**0.000**	1.000	**0.000**	**0.000**
				$\Delta BS = \Delta DEP$					
AIC	-11.75	4	2	**0.000**	0.659	**0.002**	**0.001**	**0.000**	**0.000**
SIC	-11.30	1	2	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
				$\Delta BS = \Delta BOR$					
AIC	-12,05	4	3	**0.000**	0.537	**0.000**	**0.000**	**0.000**	**0.000**
SIC	-11.57	1	3	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**

Note: The table displays the minimum value of the criteria achieved over values of $p \in [1, 4]$ and $d \in [1, 4]$ for both the linear (VAR) and threshold (TVAR) models and displays the associated values of p and d for the trheshold model. Residual autocorrelation and ARCH tests are LM tests based on 12 lags applied to the specifications described in columns 3 to 5. Only p-values for the tests on the residuals of each single equation in the VAR are reported. In bold are the test-stats rejecting the null hypothesis (at the 10% significance level) of serial correlation and heteroskedasticity.

Table 6.5: Parameter selection: Foreign banks/CP-Tbill spread

This table presents the parameters (p and d) and residual tests associated with the specifications of the TVAR chosen by the AIC and SIC criteria respectively. We consider specifications with up to 8 lags and delay up to 8 weeks. The models estimated are 3-variable TVAR models including the changes in bank balance sheet variables (ΔBS), the Tbill rate (ΔTB) and the commercial paper-Tbill spread ($CPTB$) in that order. The threshold variable is the spread variable $CPTB$. The data is weekly from 1-Jan-1990 to 1-Aug-2007. ΔTB is the first difference of the annualized 3-month Tbill rate (FRED). $CPTB$ is constructed as the 3-month AA (non-financial) commercial paper rate minus the 3-month Tbill rate (Global Financial data). Commercial bank balance sheet data is from the Federal Reserve Bank H.8 Release. All series are seasonally adjusted. On the asset side, TA stands for total assets, $LOAN$ are total loans, CIL are commercial and industrial loans, REL are real estate loans and LIQ are liquid assets (sum of securities, Fed Funds, reverse repos and cash assets). On the liability side, DEP are demand deposits, LTD are large time deposits, BOR is credit obtained from banks and others (incl. Fed Funds and repurchase agreements). Changes are normalized and expressed in percentage terms as described in Table 6.2.

				Autocorrelation tests			ARCH tests		
	TVAR	p	d	ΔBS	ΔTB	$CPTB$	ΔBS	ΔTB	$CPTB$
				$\Delta BS = \Delta LIQ$					
AIC	-10.63	4	3	**0.001**	0.609	**0.000**	**0.000**	**0.000**	**0.000**
SIC	-10.18	1	2	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
				$\Delta BS = \Delta LOAN$					
AIC	-11.40	4	2	0.974	0.658	**0.000**	0.136	**0.000**	**0.000**
SIC	-11.00	1	2	0.745	**0.000**	**0.000**	0.129	**0.000**	**0.000**
				$\Delta BS = \Delta CIL$					
AIC	-13.63	4	1	0.341	0.603	**0.000**	0.309	**0.000**	**0.000**
SIC	-13.21	1	2	**0.077**	**0.000**	**0.000**	0.522	**0.000**	**0.000**
				$\Delta BS = \Delta REL$					
AIC	-15.62	4	2	**0.000**	0.812	**0.001**	**0.000**	**0.000**	**0.000**
SIC	-15.10	1	2	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
				$\Delta BS = \Delta DEP$					
AIC	-15.51	4	2	**0.000**	0.597	**0.000**	**0.000**	**0.000**	**0.000**
SIC	-15.04	1	2	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
				$\Delta BS = \Delta LTD$					
AIC	-11.42	4	2	0.220	0.797	**0.000**	**0.000**	**0.000**	**0.000**
SIC	-11.02	1	2	0.329	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
				$\Delta BS = \Delta BOR$					
AIC	-10.39	4	4	0.687	0.684	**0.000**	**0.046**	**0.000**	**0.000**
SIC	-9.95	1	2	**0.042**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**

Note: The table displays the minimum value of the criteria achieved over values of $p \in [1, 4]$ and $d \in [1, 4]$ for both the linear (VAR) and threshold (TVAR) models and displays the associated values of p and d for the trheshold model. Residual autocorrelation and ARCH tests are LM tests based on 12 lags applied to the specifications described in columns 3 to 5. Only p-values for the tests on the residuals of each single equation in the VAR are reported. In bold are the test-stats rejecting the null hypothesis (at the 10% significance level) of serial correlation and heteroskedasticity.

Table 6.6: Parameter selection: Large banks/TED spread

This table presents the parameters (p and d) and residual tests associated with the specifications of the TVAR chosen by the AIC and SIC criteria respectively. We consider specifications with up to 4 lags and delay up to 4 weeks. The models estimated are 3-variable TVAR models including the changes in bank balance sheet variables (ΔBS), the Tbill rate (ΔTB) and the TED spread (TED) in that order. The threshold variable is the spread variable TED. The data is weekly from 1-Jan-1990 to 1-Aug-2007. ΔTB is the first difference of the annualized 3-month Tbill rate (FRED). TED is constructed as the 3-month LIBOR rate minus the 3-month Tbill rate (FRED). Commercial bank balance sheet data is from the Federal Reserve Bank H.8 Release. All series are seasonally adjusted. On the asset side, TA stands for total assets, $LOAN$ are total loans, CIL are commercial and industrial loans, REL are real estate loans and LIQ are liquid assets (sum of securities, Fed Funds, reverse repos and cash assets). On the liability side, DEP are demand deposits, LTD are large time deposits, BOR is credit obtained from banks and others (incl. Fed Funds and repurchase agreements). Changes are normalized and expressed in percentage terms as described in Table 6.2.

				Autocorrelation tests			ARCH tests		
	TVAR	p	d	ΔBS	ΔTB	TED	ΔBS	ΔTB	TED
				$\Delta BS = \Delta LIQ$					
AIC	-13.15	4	1	**0.000**	0.497	**0.005**	**0.000**	**0.000**	**0.000**
SIC	-12.71	1	1	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
				$\Delta BS = \Delta LOAN$					
AIC	-13.87	4	1	0.543	0.563	**0.010**	0.702	**0.000**	**0.000**
SIC	-13.45	1	2	**0.073**	**0.000**	**0.000**	0.981	**0.000**	**0.000**
				$\Delta BS = \Delta CIL$					
AIC	-16.64	4	1	**0.008**	0.491	**0.024**	0.674	**0.000**	**0.000**
SIC	-16.21	1	2	**0.000**	**0.000**	**0.000**	0.585	**0.000**	**0.000**
				$\Delta BS = \Delta REL$					
AIC	-15.00	4	1	0.852	0.549	**0.010**	1.000	**0.000**	**0.000**
SIC	-14.59	1	2	0.107	**0.000**	**0.000**	1.000	**0.000**	**0.000**
				$\Delta BS = \Delta DEP$					
AIC	-13.21	4	1	**0.000**	0.449	**0.015**	0.246	**0.000**	**0.000**
SIC	-12.72	1	1	**0.000**	**0.000**	**0.000**	**0.007**	**0.000**	**0.000**
				$\Delta BS = \Delta LTD$					
AIC	-16.80	4	1	**0.000**	0.656	**0.020**	**0.000**	**0.000**	**0.000**
SIC	-16.34	1	3	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
				$\Delta BS = \Delta BOR$					
AIC	-13.84	4	1	0.171	0.487	**0.002**	**0.044**	**0.000**	**0.000**
SIC	-13.42	1	1	**0.007**	**0.000**	**0.000**	**0.031**	**0.000**	**0.000**

Note: The table displays the minimum value of the criteria achieved over values of $p \in [1, 4]$ and $d \in [1, 4]$ and displays the associated values of p and d for the trheshold model. Residual autocorrelation and ARCH tests are LM tests based on 12 lags applied to the specifications described in columns 3 to 5. Only p-values for the tests on the residuals of each single equation in the VAR are reported. In bold are the test-stats rejecting the null hypothesis (at the 10% significance level) of serial correlation and heteroskedasticity.

Table 6.7: Threshold estimation and linearity tests: Large banks/CP-Tbill spread

The models estimated are 3-variable TVAR models including 1 lag of the changes in bank balance sheet variables (ΔBS), the Tbill rate (ΔTB) and the commercial paper-Tbill spread ($CPTB$) in that order. The threshold variable is the spread variable $CPTB$ with a delay of 2 weeks. The data is weekly from 4-Jan-1990 to 1-Aug-2007. Commercial bank balance sheet data is from the Federal Reserve Bank H.8 Release. All series are seasonally adjusted. On the asset side, TA stands for total assets, $LOAN$ are total loans, CIL are commercial and industrial loans, REL are real estate loans and LIQ are liquid assets (sum of securities, Fed Funds, reverse repos and cash assets). On the liability side, DEP are demand deposits, LTD are large time deposits, BOR is credit obtained from banks and others (incl. Fed Funds and repurchase agreements). Changes are normalized and expressed in percentage terms as described in Table 6.2. ΔTB is the first difference of the annualized 3-month Tbill rate (FRED). $CPTB$ is constructed as the 3-month AA (non-financial) commercial paper rate minus the 3-month Tbill rate (Global Financial data). The table displays the threshold estimates for each TVAR alongside the Sup Wald test statistics and its asymptotic p-value.

	ΔLIQ	$\Delta LOAN$	ΔCIL	ΔREL	ΔDEP	ΔLTD	ΔBOR
Panel A: Full VAR							
Threshold est.	0.30	0.26	0.53	0.26	0.26	0.15	0.26
Sup-Wald	100.59	97.22	115.92	90.01	98.39	132.61	103.85
p-value	0.008	0.050	0.126	0.006	0.01	0.000	0.002
Panel B: equation for ΔBS							
Threshold est.	0.33	0.34	0.48	0.45	0.14	0.14	0.40
Sup-Wald	9.76	6.96	11.72	8.51	8.09	21.45	8.87
p-value	0.284	0.566	0.146	0.400	0.484	0.002	0.386
Panel C: equation for ΔTB							
Threshold est.	0.27	0.27	0.27	0.27	0.22	0.27	0.27
Sup-Wald	26.64	25.48	28.38	26.65	27.06	24.06	27.80
p-value	0.002	0.002	0.000	0.002	0.000	0.002	0.000
Panel D: equation for $CPTB$							
Threshold est.	0.48	0.26	0.51	0.48	0.48	0.26	0.26
Sup-Wald	39.49	60.50	37.66	26.45	40.51	35.15	38.13
p-value	0.000	0.000	0.000	0.002	0.000	0.000	0.000

Note: The threshold estimate is the value of the threshold variable at which the log determinant of the 'structural' residuals is minimized. The Sup-Wald test is constructed using the TVAR in structural form and accounts for a general form of heteroskedasticity. Asymptotic p-values (in bracket) for the Sup-Wald test are computed according to Hansen (1996) using 500 bootstrap replications.

Table 6.8: Threshold estimation and linearity tests: Large banks/TED spread

The models estimated are 3-variable TVAR models including 1 lag of the changes in bank balance sheet variables (ΔBS), the Tbill rate (ΔTB) and the TED spread (TED) in that order. The threshold variable is the spread variable TED with a delay of 2 weeks. The data is weekly from 4-Jan-1990 to 1-Aug-2007. Commercial bank balance sheet data is from the Federal Reserve Bank H.8 Release. All series are seasonally adjusted. On the asset side, TA stands for total assets, $LOAN$ are total loans, CIL are commercial and industrial loans, REL are real estate loans and LIQ are liquid assets (sum of securities, Fed Funds, reverse repos and cash assets). On the liability side, DEP are demand deposits, LTD are large time deposits, BOR is credit obtained from banks and others (incl. Fed Funds and repurchase agreements). Changes are normalized and expressed in percentage terms as described in Table 6.2. ΔTB is the first difference of the annualized 3-month Tbill rate (FRED). TED is constructed as the 3-month LIBOR rate minus the 3-month Tbill rate (FRED). The table displays the threshold estimates for each TVAR alongside the Sup Wald test statistics and its asymptotic p-value.

	ΔLIQ	$\Delta LOAN$	ΔCIL	ΔREL	ΔDEP	ΔLTD	ΔBOR
Panel A: Full VAR							
Threshold est.	0.66	0.49	0.66	0.66	0.66	0.34	0.66
Sup-Wald	60.87	53.02	60.81	65.21	60.76	51.46	52.65
p-value	0.060	0.194	0.17	0.042	0.112	0.044	0.072
Panel B: equation for ΔBS							
Threshold est.	0.40	0.44	0.72	0.24	0.40	0.24	0.49
Sup-Wald	10.31	6.38	9.28	10.83	14.30	15.98	5.92
p-value	0.276	0.684	0.320	0.194	0.062	0.022	0.734
Panel C: equation for ΔTB							
Threshold est.	0.39	0.34	0.34	0.39	0.24	0.34	0.39
Sup-Wald	14.66	14.93	16.95	14.33	15.82	14.27	13.27
p-value	0.116	0.090	0.058	0.094	0.082	0.100	0.158
Panel D: equation for $CPTB$							
Threshold est.	0.66	0.66	0.66	0.66	0.73	0.66	0.66
Sup-Wald	47.38	32.83	26.87	34.65	38.03	28.69	30.23
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note: The threshold estimate is the value of the threshold variable at which the log determinant of the ‘structural’ residuals is minimized. The Sup-Wald test is constructed using the TVAR in structural form and accounts for a general form of heteroskedasticity. Asymptotic p-values (in bracket) for the Sup-Wald test are computed according to Hansen (1996) using 500 bootstrap replications.

Table 6.9: Threshold estimation and linearity tests: Small banks/CP-Tbill spread

The models estimated are 3-variable TVAR models including 1 lag of the changes in bank balance sheet variables (ΔBS), the Tbill rate (ΔTB) and the commercial paper-Tbill spread ($CPTB$) in that order. The threshold variable is the spread variable $CPTB$ with a delay of 2 weeks. The data is weekly from 4-Jan-1990 to 1-Aug-2007. Commercial bank balance sheet data is from the Federal Reserve Bank H.8 Release. All series are seasonally adjusted. On the asset side, TA stands for total assets, $LOAN$ are total loans, CIL are commercial and industrial loans, REL are real estate loans and LIQ are liquid assets (sum of securities, Fed Funds, reverse repos and cash assets). On the liability side, DEP are demand deposits, LTD are large time deposits, BOR is credit obtained from banks and others (incl. Fed Funds and repurchase agreements). Changes are normalized and expressed in percentage terms as described in Table 6.2. ΔTB is the first difference of the annualized 3-month Tbill rate (FRED). $CPTB$ is constructed as the 3-month AA (non-financial) commercial paper rate minus the 3-month Tbill rate (Global Financial data). The table displays the threshold estimates for each TVAR alongside the Sup Wald test statistics and its asymptotic p-value.

	ΔLIQ	$\Delta LOAN$	ΔCIL	ΔREL	ΔDEP	ΔBOR
Panel A: Full VAR						
Threshold est.	0.53	0.26	0.26	0.26	0.53	0.26
Sup-Wald	101.57	125.32	129.11	111.23	116.38	127.51
p-value	0.102	0.006	0.004	0.008	0.106	0.016
Panel B: equation for ΔBS						
Threshold est.	0.42	0.53	0.39	0.22	0.22	0.13
Sup-Wald	7.25	11.12	7.30	11.46	12.55	9.13
p-value	0.626	0.178	0.568	0.188	0.134	0.398
Panel C: equation for ΔTB						
Threshold est.	0.30	0.27	0.27	0.27	0.30	0.53
Sup-Wald	26.92	27.81	24.72	26.08	26.03	26.58
p-value	0.004	0.000	0.002	0.000	0.002	0.000
Panel D: equation for $CPTB$						
Threshold est.	0.53	0.25	0.48	0.26	0.53	0.26
Sup-Wald	55.97	61.54	53.84	31.55	64.15	63.77
p-value	0.000	0.000	0.000	0.000	0.000	0.000

Note: The threshold estimate is the value of the threshold variable at which the log determinant of the 'structural' residuals is minimized. The Sup-Wald test is constructed using the TVAR in structural form and accounts for a general form of heteroskedasticity. Asymptotic p-values (in bracket) for the Sup-Wald test are computed according to Hansen (1996) using 500 bootstrap replications.

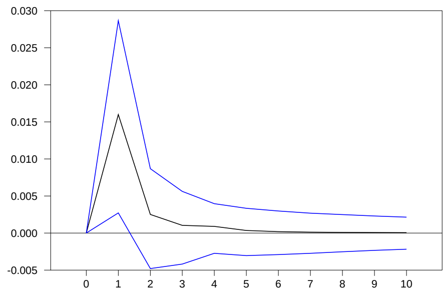
Table 6.10: Threshold estimation and linearity tests: Foreign banks/CP-Tbill spread

The models estimated are 3-variable TVAR models including 1 lag of the changes in bank balance sheet variables (ΔBS), the Tbill rate (ΔTB) and the commercial paper-Tbill spread ($CPTB$) in that order. The threshold variable is the spread variable $CPTB$ with a delay of 2 weeks. The data is weekly from 4-Jan-1990 to 1-Aug-2007. Commercial bank balance sheet data is from the Federal Reserve Bank H.8 Release. All series are seasonally adjusted. On the asset side, TA stands for total assets, $LOAN$ are total loans, CIL are commercial and industrial loans, REL are real estate loans and LIQ are liquid assets (sum of securities, Fed Funds, reverse repos and cash assets). On the liability side, DEP are demand deposits, LTD are large time deposits, BOR is credit obtained from banks and others (incl. Fed Funds and repurchase agreements). Changes are normalized and expressed in percentage terms as described in Table 6.2. ΔTB is the first difference of the annualized 3-month Tbill rate (FRED). $CPTB$ is constructed as the 3-month AA (non-financial) commercial paper rate minus the 3-month Tbill rate (Global Financial data). The table displays the threshold estimates for each TVAR alongside the Sup Wald test statistics and its asymptotic p-value.

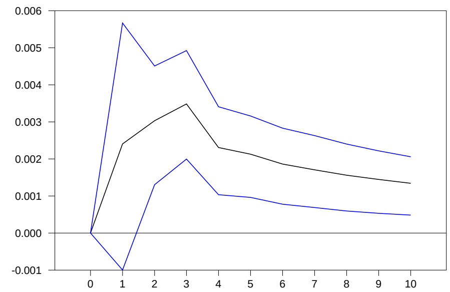
	ΔLIQ	$\Delta LOAN$	ΔCIL	ΔREL	ΔDEP	ΔLTD	ΔBOR
Panel A: Full VAR							
Threshold est.	0.26	0.26	0.26	0.48	0.26	0.26	0.26
Sup-Wald	97.66	96.65	117.04	96.51	88.94	82.26	127.16
p-value	0.002	0.012	0.010	0.056	0.032	0.022	0.010
Panel B: equation for ΔBS							
Threshold est.	0.16	0.46	0.21	0.44	0.48	0.51	0.31
Sup-Wald	14.64	3.93	25.48	16.05	8.40	10.98	7.59
p-value	0.074	0.966	0.000	0.022	0.426	0.236	0.612
Panel C: equation for ΔTB							
Threshold est.	0.27	0.27	0.26	0.27	0.27	0.27	0.27
Sup-Wald	26.63	24.41	25.65	27.91	25.65	24.04	23.57
p-value	0.002	0.004	0.002	0.000	0.002	0.002	0.006
Panel D: equation for $CPTB$							
Threshold est.	0.26	0.26	0.26	0.48	0.53	0.26	0.26
Sup-Wald	26.00	37.38	33.36	26.70	27.86	26.40	46.61
p-value	0.008	0.002	0.000	0.002	0.000	0.002	0.000

Note: The threshold estimate is the value of the threshold variable at which the log determinant of the 'structural' residuals is minimized. The Sup-Wald test is constructed using the TVAR in structural form and accounts for a general form of heteroskedasticity. Asymptotic p-values (in bracket) for the Sup-Wald test are computed according to Hansen (1996) using 500 bootstrap replications.

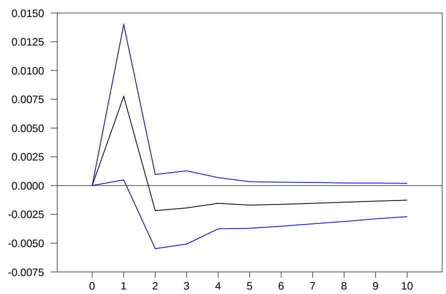
Figure 6-2: Linear VAR: Impulse responses to a CP-Tbill spread shock (large banks)



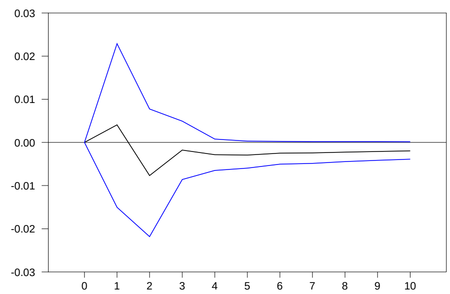
(a) Loans



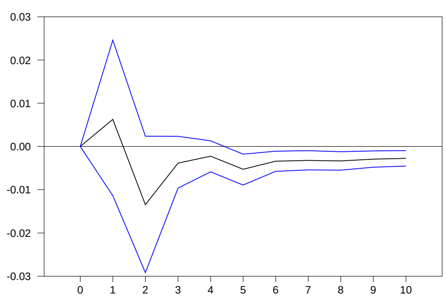
(b) Commercial and industrial loans



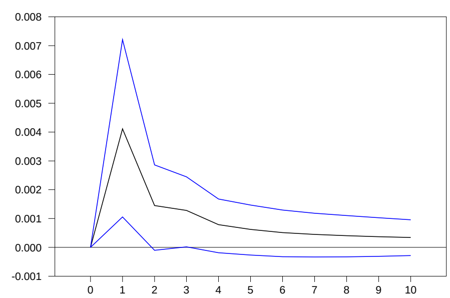
(c) Real estate loans



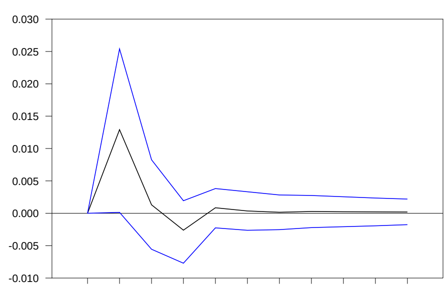
(d) Liquid assets



(e) Deposits



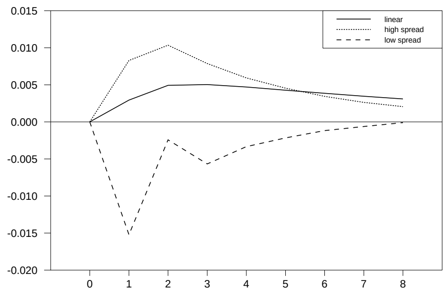
(f) Large time deposits



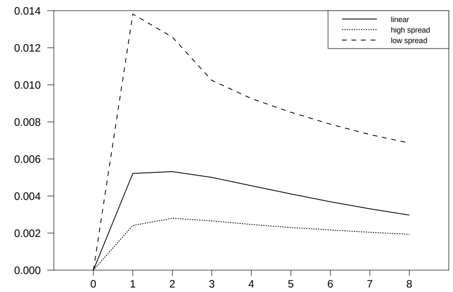
(g) Borrowing

Note: The figure displays the impulse responses of asset growth variables to a one standard deviation shock to the CP-Tbill spread from a 3-variable linear VAR model including 2 lags (as chosen by Schwartz criterion) of bank asset growth, the change in the Tbill rate and the paper-bill spread estimated over the sample period January 1990 - August 2007. Responses are in % points. The 90% confidence bands are computed using 1000 bootstrapped replications.

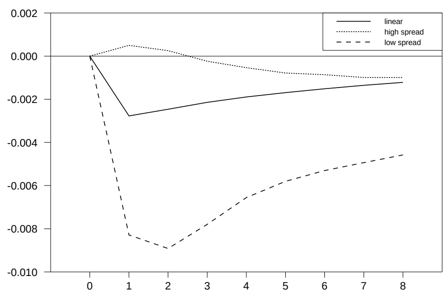
Figure 6-3: Threshold VAR: Impulse responses to a CP-Tbill spread shock (large banks)



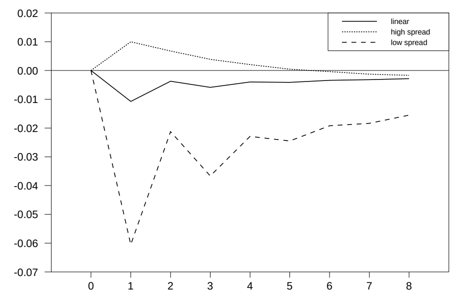
(a) Loans



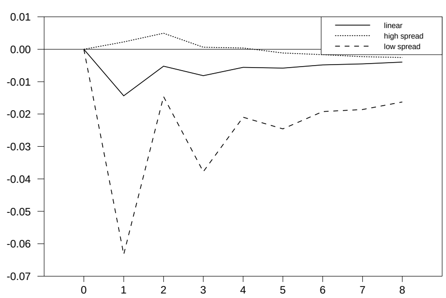
(b) Commercial and industrial loans



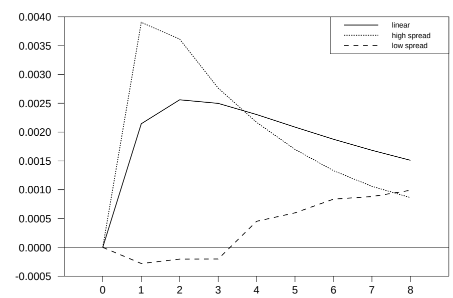
(c) Real estate loans



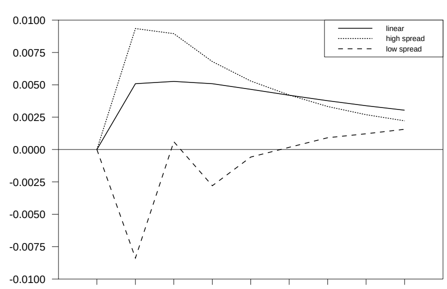
(d) Liquid assets



(e) Deposits



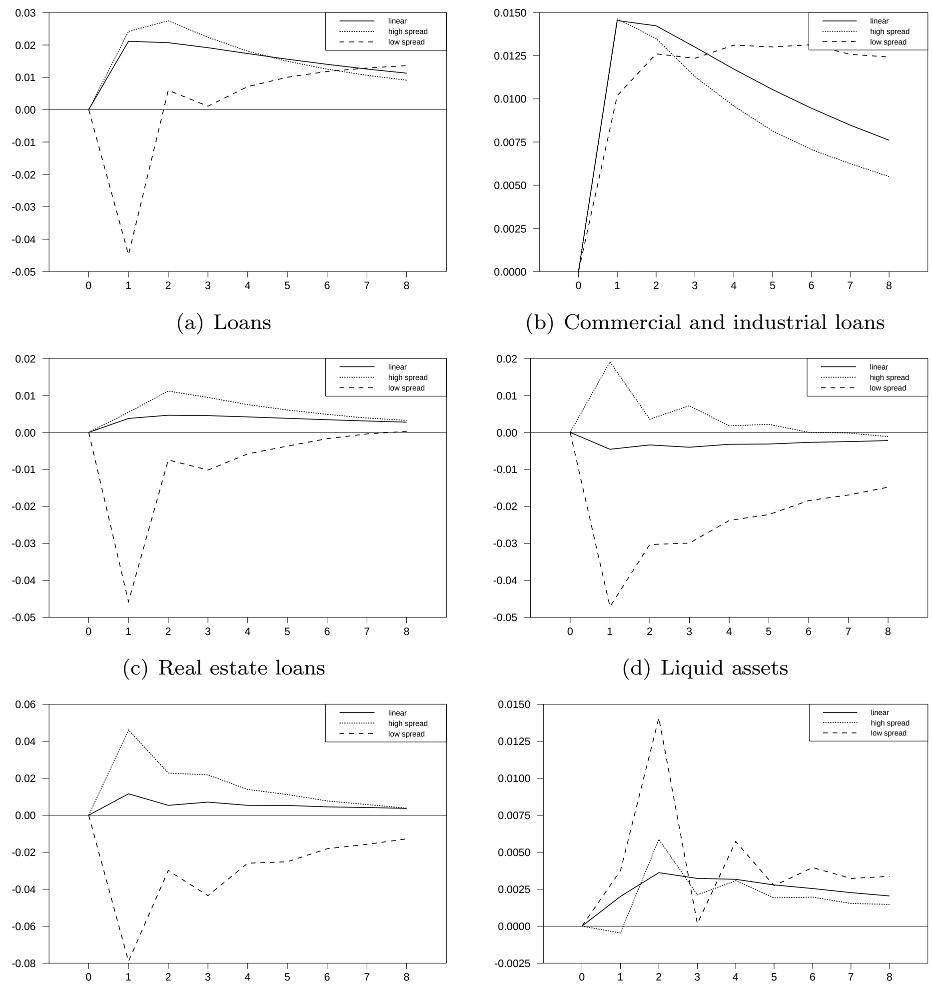
(f) Large time deposits



(g) Borrowings

Note: The figure displays the impulse responses of balance sheet variables to an 8 basis point shock to the CP-Tbill spread in both the high spread and low spread regime from a 3-variable TVAR model including 1 lag of bank asset growth, the change in the Tbill rate and the paper-bill spread. The threshold variable is the paper-bill spread with a delay of 2 weeks and the threshold estimate is set to 0.26 in all models. The impulse response from the linear model is also displayed. Responses are in % points.

Figure 6-4: Threshold VAR: Impulse responses to a CP-Tbill spread shock (small banks)



(a) Loans

(b) Commercial and industrial loans

(c) Real estate loans

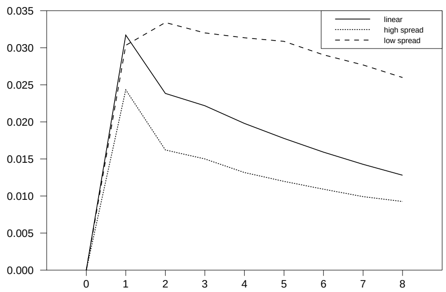
(d) Liquid assets

(e) Deposits

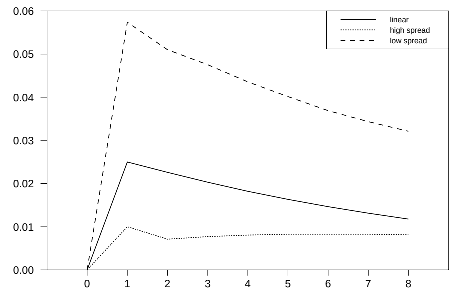
(f) Borrowings

Note: The figure displays the impulse responses of balance sheet variables to an 8 basis point shock to the CP-Tbill spread in both the high spread and low spread regime from a 3-variable TVAR model including 1 lag of bank asset growth, the change in the Tbill rate and the paper-bill spread. The threshold variable is the paper-bill spread with a delay of 2 weeks and the threshold estimate is set to 0.26 in all models. The impulse response from the linear model is also displayed. Responses are in % points.

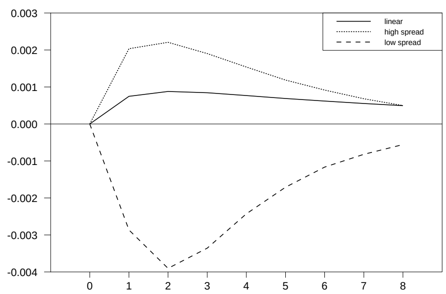
Figure 6-5: Threshold VAR: Impulse responses to a CP-Tbill spread shock (foreign banks)



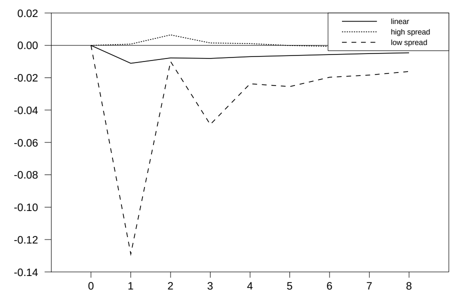
(a) Loans



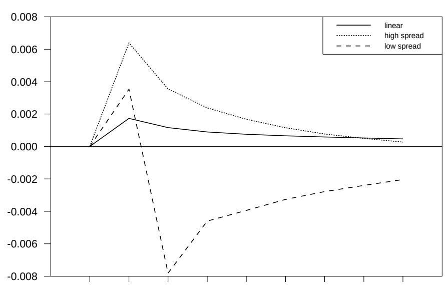
(b) Commercial and industrial loans



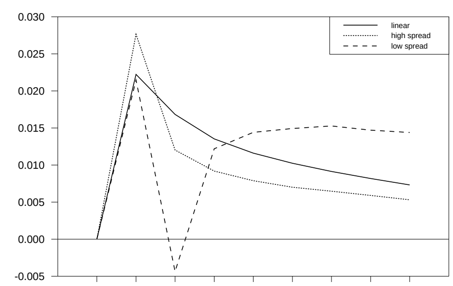
(c) Real estate loans



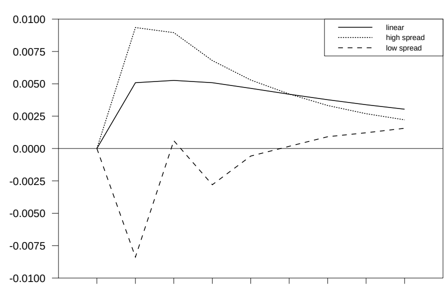
(d) Liquid assets



(e) Deposits



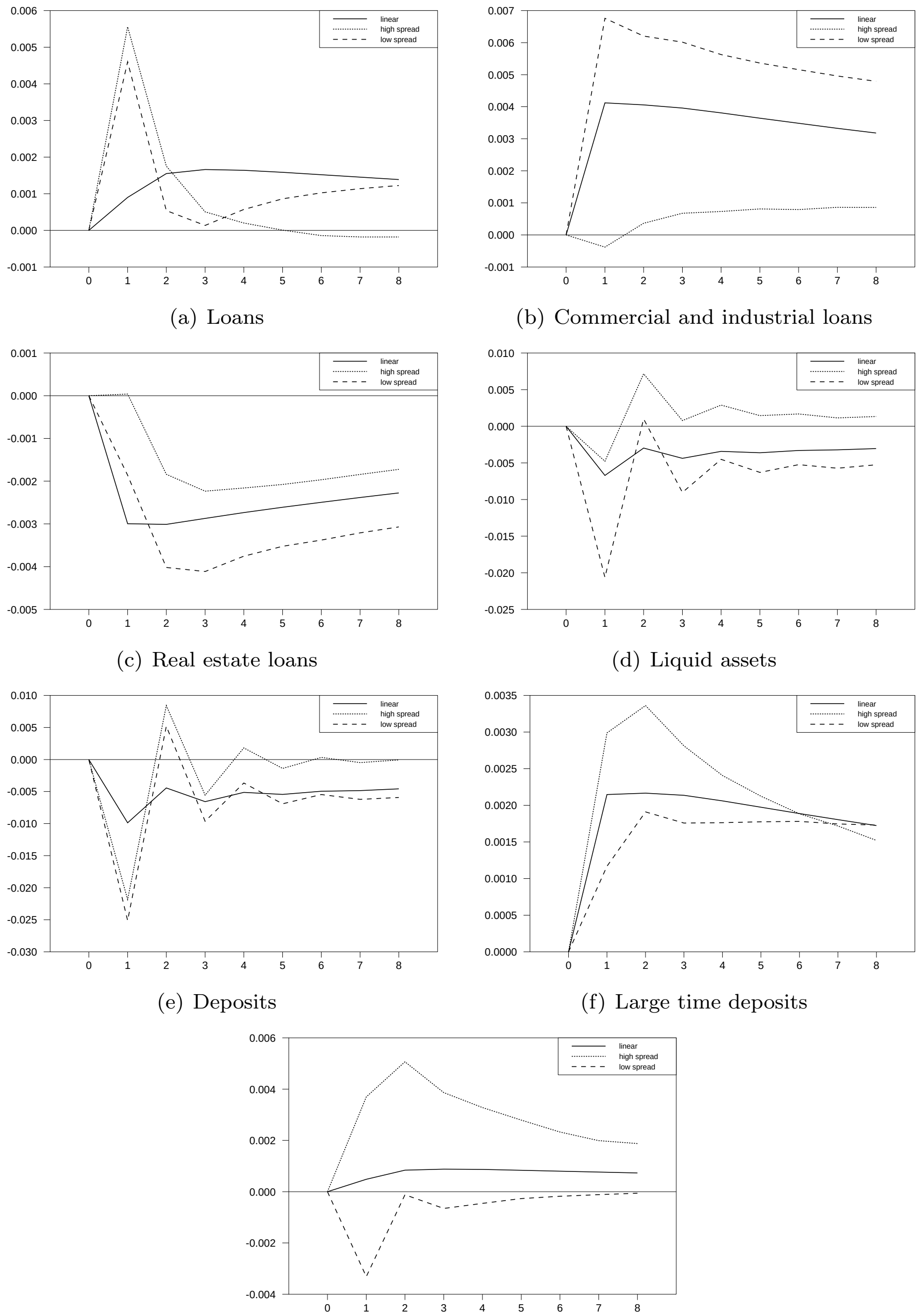
(f) Large time deposits



(g) Borrowings

Note: The figure displays the impulse responses of balance sheet variables to an 8 basis point shock to the CP-Tbill spread in both the high spread and low spread regime from a 3-variable TVAR model including 1 lag of bank asset growth, the change in the Tbill rate and the paper-bill spread. The threshold variable is the paper-bill spread with a delay of 2 weeks and the threshold estimate is set to 0.26 in all models. The impulse response from the linear model is also displayed. Responses are in % points.

Figure 6-6: Threshold VAR: Impulse responses to a TED spread shock (large banks)



(a) Loans

(b) Commercial and industrial loans

(c) Real estate loans

(d) Liquid assets

(e) Deposits

(f) Large time deposits

(g) Borrowings

Note: The figure displays the impulse responses of balance sheet variables to an 8 basis point shock to the TED spread in both the high spread and low spread regime from a 3-variable TVAR model including 1 lag of bank asset growth, the change in the Tbill rate and the TED spread. The threshold variable is the TED spread with a delay of 2 weeks and the threshold estimate is set to 0.66 in all models. The impulse response from the linear model is also displayed. Responses are in % points.

Appendix

A. Description of the analytical expressions for the Wald tests

Let T be the number of observations in the sample, k the number of variables and p the number of lags.

A1. Reduced-form model

$$Y_t = C_1 + B_1(L)Y_{t-1} + (C_2 + B_2(L)Y_{t-1})I(q_{t-d} > \gamma) + u_t \tag{6.5}$$

where q_t is the threshold variable, γ is the threshold value and d is the delay parameter. Rewriting (6.5) in a more compact form:

$$Y_t = BX_t + u_t \tag{6.6}$$

where $Y_t = [y_{1t} \; ... \; y_{kt}]'_{k\times 1}$, $B = [B_1 \;\; B_2]_{\,k\times 2(pk+1)}$ $\big(B_i = [C_i \;\; B_{i1} \; ... \; B_{ip}]_{\,k\times(pk+1)}\big)$, $Z_t = [1 \; Y'_{t-1} \; ... \; Y'_{t-p}]'_{(pk+1)\times 1}$, $X_t = [Z'_t \; d_t Z'_t]'_{2(pk+1)\times 1}$, $u_t = [u_{1t} \; ... \; u_{kt}]'_{k\times 1}$ and

$$d_t = \begin{cases} 1 & \text{if } q_{t-d} \geq \gamma \\ 0 & \text{if } q_{t-d} < \gamma \end{cases}$$

Rewriting (6.6) over the full sample:

$$Y = BX + U \tag{6.7}$$

where $Y = [Y_1 \; ... \; Y_T]'_{k\times T}$, $X = [X_1 \; ... \; X_T]'_{2(kp+1)\times T}$ and $U = [u_1 \; ... \; u_T]'_{k\times T}$. Vectorizing (6.7):

$$y = (X' \otimes I_k)\beta + u \tag{6.8}$$

where $y = \text{vec}(Y)$ is a $kT \times 1$ vector, $\beta = [\beta'_1 \;\; \beta'_2]' = \text{vec}(B)$ is a $2k(kp+1) \times 1$ vector and $u = \text{vec}(U)$ is a $kT \times 1$ vector.

The LS estimators for β and the variance covariance matrix of residuals are given by:

$$\hat{\beta} = [(XX')^{-1}X \otimes I_k]y \tag{6.9}$$

$$\hat{\Sigma}_u = \frac{1}{T-(kp+1)}[Y - \hat{B}X][Y - \hat{B}X]' \quad \text{where} \quad \text{vec}(\hat{B}) = \hat{\beta} \tag{6.10}$$

The variance of the LS estimator is:

$$\text{var}(\hat{\beta})^{homo} = (XX')^{-1} \otimes \hat{\Sigma}_u \tag{6.11}$$

and a heteroskedasticity consistent estimate of the variance of $\hat{\beta}$ is:

$$\text{var}(\hat{\beta})^{hetero} = [(XX')^{-1} \otimes I_k] \left(\sum_{t=1}^{T} (X_t X_t' \otimes \Sigma_{u_t}) \right) [(XX')^{-1} \otimes I_k] \tag{6.12}$$

Testing for a threshold effect is equivalent to testing the null hypothesis of a linear VAR $H_0 : \beta_2 = 0$, which can be rewritten as $H_0 : L\beta = 0$ where $L = [0_{k(kp+1)\times k(kp+1)} \;\; I_{k(kp+1)}]$, against the 2-regime TVAR alternative. The Wald test for each possible $(\gamma, d) \in (\Gamma, D)$ is formed as follows:

$$W(\gamma, d) = (L\hat{\beta})'[L\text{var}(\hat{\beta})L']^{-1}(L\hat{\beta}) \tag{6.13}$$

A.2. Structural-form model

$$Y_t = C_1 + A_1 Y_t + B_1(L)Y_{t-1} + (C_2 + A_2 Y_t + B_2(L)Y_{t-1})I(q_{t-d} > \gamma) + \epsilon_t$$

where q_t is the threshold variable, γ is the threshold value and d is the delay parameter. Rewriting (6.1) in a more compact form:

$$Y_t = BX_t + \epsilon_t \tag{6.14}$$

where $Y_t = [y_{1t} \; ... \; y_{kt}]'_{k\times 1}$, $B = [B_1 \; B_2]_{\; k\times 2[k(p+1)+1]}$ $(B_i = [C_i \; A_i \; B_{i1} \; ... \; B_{ip}]_{\; k\times[k(p+1)+1]})$, $Z_t = [1 \; Y_t' \; Y_{t-1}' \; ... \; Y_{t-p}']'_{[k(p+1)+1]\times 1}$, $X_t = [Z_t' \; d_t Z_t']'_{2[k(p+1)+1]\times 1}$, $\epsilon_t = [\epsilon_{1t} \; ... \; \epsilon_{kt}]'_{k\times 1}$ and

$$d_t = \begin{cases} 1 & \text{if } q_{t-d} \geq \gamma \\ 0 & \text{if } q_{t-d} < \gamma \end{cases}$$

Rewriting (6.14) over the full sample:

$$Y = BX + E \tag{6.15}$$

where $Y = [Y_1 \; ... \; Y_T]'_{k\times T}$, $X = [X_1 \; ... \; X_T]'_{2[k(p+1)+1]\times T}$ and $E = [\epsilon_1 \; ... \; \epsilon_T]'_{k\times T}$. Vectorizing (6.15):

$$y = (X' \otimes I_k)\beta + e \tag{6.16}$$

where $y = \text{vec}(Y)$ is a $kT \times 1$ vector, $\beta = [\beta_1' \; \beta_2']' = \text{vec}(B)$ is a $2k[k(p+1)+1] \times 1$ vector and $e = \text{vec}(E)$ is a $kT \times 1$ vector.

Next we impose the recursive structure on the β vector ($(k^2 + k)/2$ zero restrictions in each A_i matrix), which is equivalent to estimating it subject to linear constraints of the form:

$$\beta = A\lambda \tag{6.17}$$

where A is the $2k[k(p+1)+1] \times R$ matrix of restrictions (where $R = 2k[k(p+1)+1] - (k^2+k)$ is the number of unrestricted parameters) and $\lambda = [\lambda_1' \; \lambda_2']'$ is the $R \times 1$ vector of parameters to be estimated. Substituting the constraint (6.17) in (6.16) gives:

$$y = (X' \otimes I_k)A\lambda + e \tag{6.18}$$

Since in a recursive system the error terms are uncorrelated, the parameters can be estimated by OLS. The LS estimators for λ and the variance covariance matrix of residuals are given by:

$$\hat{\lambda} = [A'((XX')^{-1} \otimes I_k)A]^{-1}A'(X \otimes I_k)y \tag{6.19}$$

$$\hat{\Sigma}_e = \frac{1}{T - [k(p+1)+1]}[Y - \hat{B}X][Y - \hat{B}X]' \quad \text{where} \quad \text{vec}(\hat{B}) = \hat{\beta}, \quad \hat{\beta} = A\hat{\lambda} \tag{6.20}$$

where $k(p+1)+1$ is the maximum number of parameters estimated in any equation of the system (the last one in a recursive system).
The variance of the LS estimator is:

$$\text{var}(\hat{\lambda})^{homo} = [A'((XX')^{-1} \otimes I_k)A]^{-1}A'((XX')^{-1} \otimes \hat{\Sigma}_e)A[A'((XX')^{-1} \otimes I_k)A]^{-1} \tag{6.21}$$

and a heteroskedasticity consistent estimate of the variance of $\hat{\lambda}$ is:

$$\text{var}(\hat{\lambda})^{hetero} = [A'((XX')^{-1} \otimes I_k)A]^{-1}A'\left(\sum_{t=1}^{T}(X_tX_t' \otimes \Sigma_{u_t})\right)A[A'((XX')^{-1} \otimes I_k)A]^{-1} \tag{6.22}$$

Testing for a threshold effect is equivalent to testing the null hypothesis of a linear VAR $H_0 : \lambda_2 = 0$, which can be rewritten as $H_0 : Q\lambda = 0$ where $Q = [0_{R/2\times R/2}\ I_{R/2}]$, against the 2-regime TVAR alternative. The Wald test for each possible $(\gamma, d) \in (\Gamma, D)$ is formed as follows:

$$W(\gamma, d) = (Q\hat{\lambda})'[Q\text{var}(\hat{\lambda})Q']^{-1}(Q\hat{\lambda}) \tag{6.23}$$

B. Algorithms for the computation of asymptotic p-values (Hansen, 1996) and bootstrap p-values (Hansen, 1999)

B1. Asymptotic distribution of the sup/ave/exp Wald tests

If error terms are assumed to be conditionally homoskedastic:

1. Generate a series of multivariate Gaussian disturbances $v_t \sim \mathrm{N}(0, \Sigma^*)$. Where Σ^* is the estimated variance covariance matrix of the residuals of the TVAR model with $(\hat{\gamma}, \hat{d})$.
   - Generate a series of multivariate Gaussian disturbances $z_t \sim \mathrm{N}(0, I_k)$.
   - Compute the matrix P such that $PP' = \Sigma^*$.
   - Construct the vector series $V = PZ$ and let $v = \mathrm{vec}(V)$.
2. For each $(\gamma, d) \in (\Gamma \times D)$, set:

$$W(\gamma,d)^j_{RF} = v'((XX')^{-1}X \otimes I_k)'L'[L\mathrm{var}(\hat{\beta})^{homo}L']^{-1}L((XX')^{-1}X \otimes I_k)v$$
$$W(\gamma,d)^j_{SF} = v'G'Q'[Q\mathrm{var}(\hat{\lambda})^{homo}Q']^{-1}QGv, \quad G = [A'((XX')^{-1} \otimes I_k)A]^{-1}A'(X \otimes I_k)$$

3. Set SupW^j, AveW^j and ExpW^j.
4. Repeat 1 to 3 $J = 1000$ times and construct asymptotic p-values as:

$$p = \frac{1}{1000}\sum_{j=1}^{1000}(\mathrm{Sup}W^j \geq \mathrm{Sup}W)$$

   The p-values for the AveW and ExpW tests are computed in a similar fashion.

If conditional homoskedasticity cannot be assumed:

1. Generate a series of multivariate Gaussian disturbances $z_t \sim \mathrm{N}(0, I_k)$ and set $V_t = Z_t \odot \hat{u}^*_t$ $(V_t = Z_t \odot \hat{\epsilon}^*_t)$ where $\hat{u}^*_t$ $(\hat{\epsilon}^*_t)$ are the estimated residuals from the TVAR model with $(\hat{\gamma}, \hat{d})$. Let $v = \mathrm{vec}(V)$.
2. For each $(\gamma, d) \in (\Gamma \times D)$, set:

$$W(\gamma,d)^j_{RF} = v'((XX')^{-1}X \otimes I_k)'L'[L\mathrm{var}(\hat{\beta})^{hetero}L']^{-1}L((XX')^{-1}X \otimes I_k)v$$
$$W(\gamma,d)^j_{SF} = v'G'Q'[Q\mathrm{var}(\hat{\lambda})^{hetero}Q']^{-1}QGv, \quad G = [A'((XX')^{-1} \otimes I_k)A]^{-1}A'(X \otimes I_k)$$

3. Set SupW^j, AveW^j and ExpW^j.
4. Repeat 1 to 3 $J = 1000$ times and construct asymptotic p-values as:

$$p = \frac{1}{1000}\sum_{j=1}^{1000}(\mathrm{Sup}W^j \geq \mathrm{Sup}W)$$

C. Algorithm for the computation of GIRFs (Koop, Pesaran and Potter, 1996)

To estimate the conditional impulse response function to a structural innovation to variable i over horizon h given regime R:

1. Compute the regime specific coefficient estimates using the estimated reduced form model 6.5:

$$Y_t = (C_1 + B_1(L)Y_{t-1} + u_{1t})I(q_{t-d} \leq \gamma) + (C_2 + B_2(L)Y_{t-1} + u_{2t})I(q_{t-d} > \gamma)$$

   Compute the estimates of the variance covariance matrices in each regime $\widehat{\Sigma_1}$ and $\widehat{\Sigma_2}$. Let P_1 and P_2 be their respective Choleski decompositions. Create a series of "structural" residuals from the estimated residual series as follows:

$$\epsilon_t = \begin{cases} P_1^{-1}u_{1t} & \text{if } q_{t-d} \leq \gamma \\ P_2^{-1}u_{2t} & \text{if } q_{t-d} > \gamma \end{cases}$$

   such that $\Sigma_\epsilon = I$.

2. Pick one initial condition Y_{t-1} in regime R.

3. Create a bootstrap series of the structural innovations constructed in step 1, $\{\epsilon_t^0\}_{t=0}^h$.

4. Simulate a benchmark series $\{Y_t^0\}_{t=0}^h$ using the coefficients estimated in step 1, the initial condition from step 2 and the bootstrap series from step 3 as follows:

$$Y_t^0 = (C_1 + B_1(L)Y_{t-1} + P_1\epsilon_t^0)I(q_{t-d} \leq \gamma) + (C_2 + B_2(L)Y_{t-1} + P_2\epsilon_t^0)I(q_{t-d} > \gamma)$$

5. Simulate a shock series $\{Y_t^*\}_{t=0}^h$ using the same methodology as in step 4 but replacing the innovation to variable i in period 0, ϵ_{i0}^0, by $\eta \times \sigma_i$ where $\eta = \{1, -1, 2, -2\}$ and σ_i is the standard deviation of structural innovations to variable i. In order to compare the GIRFs across regimes σ_i is set equal to the standard deviation of the "structural" innovation to variable i estimated using the linear model.

6. Calculate the difference between the shock series $\{Y_t^*\}_{t=0}^h$ and the benchmark series $\{Y_t^0\}_{t=0}^h$.

7. Repeat step 3 to 6 B = 500 times. This yields an estimate of the conditional distribution of the GIRF given Y_{t-1} and shock η.

8. Repeat step 2 to 7 for each possible initial condition in regime R. This yields an estimate of the conditional distribution of the GIRF given R and shock η.

9. We plot the mean of this distribution calculated as follows:

$$\widehat{GIR_Y}(h, \eta, R) = \frac{1}{N_R} \sum_{j=1}^{N_R} \left\{ \frac{1}{B} \sum_{i=1}^{B} [Y_h^{*i}(\eta, Y_{h-1}^j) - Y_h^{0i}(Y_{h-1}^j)] \right\}$$

where N_R is the number of histories (initial conditions) in regime R and B is the number of bootstrap replications performed.

Chapter 7

Conclusion

In conclusion, since the main results and contributions have already been discussed in the introduction and at the end of the chapters, we will discuss some areas of future research.

In chapter 5, we investigated how changes in the Fed Funds rate affected bank lending standards in the US in the period preceding the Great Recession. Our results are consistent with the existence of a risk-taking channel. However, using a strutural VAR analysis, standard identification procedures (such as the Cholesky decomposition) may appear too restrictive to clearly identify bank risk-taking shocks. The model developed by Dell'Ariccia et. al. (2014) provides a rationale for a structural identification of such shocks. The risk-taking channel of monetary policy corresponds to a situation where an expansionary monetary policy leads to both an expansion of credit and a loosening of lending standards. We must therefore also be able to identify situations in which an expansionary monetary policy leads simultaneously to an expansion of credit and a tightening of lending standards. A possible avenue for future research will therefore be to identify bank risk-taking shocks using both loan volumes and bank lending standards. Identification schemes using sign restrictions pioneered by Uhlig (2005) represent an appropriate alternative to pursue this goal. Peersman and Wagner (2014) use this technique to identify bank risk-taking shocks in the data but identify bank risk-taking to securitisation activities.

In chapter 6, we investigated how the conditions in the markets for short-term funds had an impact on the ability of commercial banks to buffer credit supply from liquidity shocks. We find evidence in favour of the hypothesis that commercial banks are more likely to offer a better hedge against these shocks when spreads in the commercial paper market are high. This points to the existence of a source of fragility of commercial banks when the macroeconomic environment is benign as it was the case in the three years leading up to the 2007 banking crisis, a period characterized by low policy rates. A possible avenue of future research would examine how these mechanisms interact with the transmission of monetary policy. In particular, there are two ways in which policy rates may play a role. First, by affecting the aggregate supply of liquidity, the level of the Fed Funds rate may determine whether the economy is in the low-spread regime. Second, by affecting funding conditions, changes in the Fed Funds rate may lead to similar asymmetric responses in the two regimes.

Bibliography

[1] Acharya, V and Naqvi, H. (2012): "The seeds of a crisis: a theory of bank liquidity and risk taking over the business cycle", *Journal of Financial Economics*, 106, pp. 349-366.

[2] Adrian, T., Estrella, A. and Shin, H.S. (2010): "Monetary cycles, financial cycles, and the business cycle", *Federal Reserve Bank of New York Staff Reports* , Staff Report no. 421.

[3] Adrian, T. and Shin, H.S. (2011): "Financial intermediaries and monetary economics", in: B.M. Friedman and M. Woodford (eds), *Handbook of Monetary Economics 3A*, Elsevier, New York, NY.

[4] Adrian, T. and Shin, H.S. (2010): "Liquidity and leverage", *Journal of Financial Intermediation*, 19, pp 418-437.

[5] Adrian, T. and Shin, H.S. (2014); "Procyclical leverage and Value-at-Risk". *The Review of Financial Studies*, 27(2), pp. 373-403.

[6] Adrian, T., Etula, E. and Muir, T. (2014): "Financial intermediaries and the cross-section of asset returns", *Journal of Finance*, forthcoming.

[7] Angeloni, I., Faia, E. and Lo Duca, M. (2015): "Monetary policy and risk-taking", *Journal of Economic Dynamic and Control*, 52, pp. 285-307.

[8] Altissimo, F. and Violante, G.L. (2001): "The nonlinear dynamics of output and unemployment in the U.S.", *Journal of Applied Economics*, 16, pp 461-486.

[9] Atanasova, C. (2003): "Credit market imperfections and business cycle dynamics: a nonlinear approach", *Studies in Nonlinear Dynamics & Econometrics*, 7(4), pp 0-20.

[10] Ball, L. and Romer, D. (1990): "Real rigidities and the non-neutrality of money", *The Review of Economic Studies*, vol. 57, no. 2, pp. 183-203.

[11] Balke, N. S. (2000): "Credit and economic activity: credit regimes and nonlinear propagation of shocks", *The Review of Economics and Statistics*, 82(2), pp 344-349.

[12] Bekaert, G., Hoerova, M. and Lo Duca (2013): "Risk, uncertainty and monetary policy", *Journal of Monetary Economics*, 60, pp 771-788.

[13] Bernanke, B. S. and Blinder, A. S. (1988): "Credit, money and aggregate demand", *AEA Papers and Proceedings*, 78(2), pp 435-439.

[14] Bernanke, B. S. and Blinder, A. S. (1992): "The Federal Funds Rate and the channels of monetary transmission", *American Economic Review*, 82(4), pp 901-921.

[15] Bernanke, B. and Gertler, M. (1989): "Agency costs, net worth and business cycle fluctuations", *American Economic Review*, vol. 79, no. 1, pp. 14-31.

[16] Bernanke, B., Gertler, M. and Gilchrist, S. (1999): "The financial accelerator in a quantitative business cycle framework", in Taylor, J., Woodford, M. (eds), *Handbook of Macroeconomics*, vol. 1, Amsterdam, Netherlands: Elsevier, pp. 1341-1393.

[17] Borio, C. and Zhu, H. (2012): "Capital regulation, risk-taking and monetary policy: a missing link in the transmission mechanism?", *Journal of Financial Stability*, 8, pp 236-251.

[18] Bruno, V. and Shin, H.S. (2015): "Capital flows and the risk-taking channel of monetary policy", *Journal of Monetary Economics*, forthcoming.

[19] Buch, C.M., Eickmeier, S. and Prieto, E. (2014): "In search for yield? Survey-based evidence on bank risk-taking", *Journal of Economic Dynamics & Control*, 43, pp 12-30.

[20] Carlstrom, C.T. and Fuerst, T.S. (1997): "Agency costs, net worth and business cycle fluctuations: a computable general equilibrium analysis", *American Economic Review*,vol. 87, no. 5, pp. 893-910.

[21] Christiano, L.J., Eichenbaum, M., Evans, C.L. (1999): "Monetary policy shocks: what have we learned and to what", in: Taylor, J.B., Woodford, M. (Eds.), Handbook of Macroeconomics. North-Holland, Amsterdam.

[22] Cornett, M. M., McNutt, J. J., Strahan, P. E. and Tehrenian H. (2011): "Liquidity risk management and credit supply in the financial crisis", *Journal of Financial Economics*, 101, pp 297-312.

[23] Cooley, T.F., (1997): "Calibrated models", *Oxford Review of Economic Policy*, Vol. 13, No.3, pp 55-69.

[24] Dell'Ariccia, G., Igan, D. and Laeven, L. (2012): "Credit booms and lending standards: evidence from the subprime mortgage market", *Journal of Money, Credit and Banking*, 44(2-3), pp 367-384.

[25] Dell'Ariccia, G., Laeven, L and Marquez, R. (2014): "Real interest rates, leverage, and bank risk-taking", —Journal of Economic Theory, 149, pp. 65-99.

[26] Dell'Ariccia, G., Laeven, L. and Suarez, G. (2013). "Bank leverage and monetary policy's risk-taking channel: evidence from the United States", *IMF Working Paper*, WP/13/143.

[27] Dell'Ariccia, G., Laeven, L. and Marquez, R. (2014): "Real interest rates, leverage, and bank risk-taking", *Journal of Economic Theory*, 149, pp 65-99.

[28] Den Haan, W. J., Sumner, S. W. and Yamashiro, G. M. (2007): "Bank loan portfolios and the monetary transmission mechanism", *Journal of Monetary Economics*, 54, pp 904-924.

[29] Freixas, X. and Rochet, J.-C. (2008). *Microeconomics of Banking*, 2^{nd} edition, Cambridge (Massachussets): The MIT Press.

[30] Gatev, E. and Strahan, P. E. (2006): "Bank's advantage in hedging liquidity risk: theory and evidence from the commercial paper market", 61(2), pp. 867-892.

[31] Gatev, E., Schuermann, T. and Strahan, E. P. (2009): "Managing bank liquidity risk: how deposit-loan synergies vary with market conditions", *Review of Financial Studies*, 22(3), pp 995-1020.

[32] Gertler, M. and Karadi, P. (2011): "A model of unconventional monetary policy", *Journal of Monetary Economics*, vol. 58, pp. 17-34.

[33] Gertler, M and Kiyotaki, K. (2011): "Financial intermediation and credit policy in business cycle analysis", in Woodford, M. and Friedmann, B. (eds), *Handbook of Monetary Economics*, Vol.3A, pp. 547-599, Amsterdam, North Holland: Elsevier.

[34] Hamilton, J. D. (1994): *Time Series Analysis*, Princeton, New Jersey: Princeton University Press.

[35] Hansen, B. E. (1996): "Inference when a nuisance parameter is not identified under the null hypothesis", *Econometrica*, 64(2), pp 413-430.

[36] Hansen, B. E. (1997): "Inference in TAR models", *Studies in Nonlinear Dynamics & Econometrics*, 2(1), pp 1-14.

[37] Hansen, B. E. (1999): "Testing for linearity", *Journal of Economic Surveys*, 13(5), pp 551-576.

[38] Hansen, B. E. (2000): "Sample splitting and threshold estimation", *Econometrica*, 68(3), pp 575-603.

[39] Holmstrom, B. and Tirole, J. (1997): "Financial intermediation, loanable funds, and the real sector", *The Quarterly Journal of Economics*, vol. 112, no. 3, pp. 663-691.

[40] Hubrich, K. and Terasvirta, T. (2013): "Threshold and smooth transitions in vector autoregressive models", *CREATES Research paper* 2013-18.

[41] Ioannidou, V., Ongena, S. and Peydro, J.-L. (2015): "Monetary policy, risk-taking and pricing: evidence form a quasi-natural experiment", *Review of Finance*, 19(1), pp 95-144.

[42] Ivashina, V. and Scharfstein, D. (2010): "Bank lending during the financial crisis of 2008", *Journal of Financial Economics*, 97, pp 319-338.

[43] Jimenez, G., Ongena, S., Peydro, J.-L. and Saurina, J. (2012): "Credit supply and monetary policy: identifying the bank balance sheet channel with loan applications", *American Economic Review*, vol. 102, no. 5, pp. 2301-2326.

[44] Jimenez, G., Ongena, S., Peydro, J.-L. and Saurina, J. (2014): "Hazardous times for monetary policy: what do twenty-three million bank loans sat about the effect of monetary policy on credit risk-taking?", *Econometrica*, 82(2), pp 463-505.

[45] Kamber, G. and Thoenissen, C. (2012): "The financial accelerator and monetary policy rules", *Economic Letters*, 115, pp. 309-313.

[46] Kashyap, A. K., Rajan, R. and Stein, J. C. (2002): "Banks as liquidity providers: an explanation for the coexistence of lending and deposit-taking", *The Journal of Finance*, 57(1), pp. 33-73.

[47] Kashyap, A. K., Stein, J. C. and Wilcox, D. W. (1993): "Monetary policy and credit conditions: evidence from the composition of external finance", *American Economic Review*, 83(1), pp. 78-98.

[48] Kashyap, A. K., Stein, J. C. and Wilcox, D. W. (1993): "Monetary policy and credit conditions: evidence from the composition of external finance: reply", *American Economic Review*, 86(1), pp. 310-314.

[49] Kiyotaki, N. and Moore, J. (1997): "Credit cycles", *Journal of Political Economy*, vol. 105, pp. 211-248.

[50] Koop, G., Pesaran, M. H. and Potter, S. M. (1996): "Impulse response analysis in nonlinear multivariate models", *Journal of Econometrics*, 74, pp 119-147.

[51] Kydland, F.E., Prescott, E.C. (1991): "The econometrics of the general equilibrium approach to business cycles", *The Scandinavian Journal of Economics*, Vol. 93, No. 2, Proceedings of a Conference on New Approaches to Empirical Macroeconomics, pp 161-178.

[52] Lown, C., Morgan, D. P. (2006): "The credit cycle and the business cycle: new findings using the loan officer opinion survey", *Journal of Money, Credit and Banking*, 38(6), pp 1555-1597.

[53] Lutkepohl, H. (2006): *New Introduction to Multiple Time Series Analysis*, Berlin: Springer.

[54] Maddaloni, A. and Peydro, J.-L. (2011): "Bank risk-taking, securitization, supervision, and low interest rates: evidence from the Euro-area and the US lending standards", *Review of Financial Studies*, 24(6), pp. 2121-2165.

[55] Miranda Agrippino, S. and Rey, H. (2014): "World asset markets and the global financial cycle", working paper.

[56] Oliner, S. D. and Rudebusch, G. D. (1996): "Monetary policy and credit conditions: evidence from the composition of external finance: comment", *American Economic Review*, 86(1), pp. 300-309.

[57] Peersman, G. and Wagner, W. (2014): "Shocks to bank lending, risk-taking, securitization, and their role for US business cycle fluctuations", WP.

[58] Pitarakis, J.-Y. (2006): "Model selection uncertainty and detection of threshold effects", *Studies in Nonlinear Dynamics & Econometrics*, 10(1), pp 0-28.

[59] Pozsar, Z., Adrian, T., Ashcraft, A. and Boesky, H. (2010): "Shadow banking", *Federal Reserve Bank of New York Staff Reports*, Staff report, no. 458.

[60] Shin, H.S. (2010): *Risk and liquidity*, Oxford: Oxford University Press.

[61] Rajan, R. G. (2005): "Has financial development made the world riskier?", *NBER Working paper series*, Working paper 11728.

[62] Tsay, R. S. (1998): "Testing and modelling multivariate threshold models", *Journal of the American Statistical Association*, 93(443), pp 1188-1202.

[63] Uhlig, H. (1995): "A toolkit for analysing nonlinear dynamic stochastic models easily", Discussion Paper 1995-97, Tilburg University, Center for Economic Research.

[64] Uhlig, H. (2005): "What are the effects of monetary policy on output? Results from an agnostic identification procedure", *Journal of Monetary Economics*, 52, pp. 381-419.